ABSTRACT

SALING, NONA E. An Empirical Study Comparing the Effect of Feedback, Training, and Executive Coaching on Leadership Behavior Change. (Under the direction of Paula Berardinelli and Brad Mehlenbacher.)

The purpose of this study was to examine the differences in leadership behavior change of participants in feedback only (A), feedback and training (B), and feedback, training, and coaching (C) groups. Participants were 101 managers, supervisors, and team leaders from businesses and nonprofit organizations in the United States. Participants received 360-degree feedback twice, based on a 28-item feedback survey. Pre-test and post-test responses were subjected to an exploratory factor analysis, resulting in two factors: relationship (14 items) and task (7 items). Seven items were discarded from further analysis. These were analyzed separately in all statistical tests of difference. There were no significant differences between the change in relationship or task factor scores over time as measured by the interaction effect of group and test, when adjusted for the covariates of time and participant age. For the task factor, effect tests found highly significant differences in both pre-test and post-test scores. Tests of pre-test scores on the task factor showed a significant difference between the feedback only (A) group and the feedback, training, and coaching (C) group. Tests of post-test scores showed a trend toward a significant difference between the feedback only (A) group and the feedback and training (B) group. Preexisting differences among the training, feedback, and coaching (C) group on the task factor may have accounted for the differences found. Based on prior research, this group may be more at risk for

derailment than the other two groups. Further research using exploratory factor analysis to identify the actual factors present in 360-degree feedback is recommended. Tests for differences between pre-tests and post-tests within each of the three groups showed trends that should be replicated in future research. Due to the small sample size of the study, results need to be interpreted with caution.

AN EMPIRICAL STUDY COMPARING THE EFFECT OF FEEDBACK, TRAINING, AND EXECUTIVE COACHING ON LEADERSHIP BEHAVIOR CHANGE

by
Nona E. Saling

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Education

Adult and Community College Education

Raleigh

2005

APPROVED BY

Paula Berardinelli	Brad Mehlenbacher
Co-chair of Advisory Committee	Co-chair of Advisory Committee
S. Bartholomew Craig	Tony O'Driscoll
Christie Knittel Mabry	

Biography

Nona E. Saling is Assistant Professor of Human Resources in the Organizational Studies Division of Peace College in Raleigh, North Carolina. She holds a B.A. in English from the University of Delaware and an M.A. in Communication from the University of Kentucky. She has held a number of management positions in training, leadership, and organization development in consumer products, financial services, publishing, and health care organizations.

Acknowledgements

First, I would like to thank my immediate family—Pat Saling, Keith Burridge, and Claire Burridge—for the support and patience they have given me during this journey that sometimes seemed without end. They have had the good grace to put up with last-minute cancellations and short tempers, and to actually pitch in and help with some of the work! I would also like to thank my students at Peace College and two of my colleagues in the Human Resources department, Heather Lee and Kathy Corley, for their enthusiastic confidence in my abilities, even when my doubt was strongest.

Dr. Traci Clayton had enough faith in me to agree to sponsor this research project. Jerry Baumgartner and Faye Atkinson have played vital roles in allowing me to carry it out. I could not have completed the project without them. Many others at Farr Associates have supported me during this project: Mark Ernsberger, Kevin Haggerty, Jim Mooney, Cary Root, Steve Royal, Cindy Jones, Jordan Neely-Chase, and Annie Sparks Baumgartner. I thank all of them for welcoming me into their professional family.

My committee has been highly supportive and very responsive to questions and requests for advice. I'd like to especially recognize Dr. Bart Craig for guiding me through the exploratory factor analysis process. Joy Smith of the NCSU Department of Statistics was also a steady guide through the sometimes murky waters of the SAS system. Fellow graduate students Lisa Wieland Handy, Varnell McDonald-Fletcher, and Jeff Jones were pillars of strength and comfort as we worked through the proposal process together.

Finally, I must thank my companion animal, Maggie, for her incredibly accurate sense of time. She unfailingly reminded me when it was 5 o'clock on days when I was

working on the dissertation. Who knows when our nutritional requirements might have been met without her very faithful reminders?

I do regret that my mother and father are not here to celebrate this accomplishment. They would have been very proud of me, and I am very grateful that they instilled in me the self-confidence and perseverance necessary to complete the doctorate.

TABLE OF CONTENTS

	Page
List of Tables	viii
List of Figures	x
Chapter 1: Introduction	1
Background: Rapid Growth	1
Definition of Executive Coaching	2
Purposes of Executive Coaching	3
The Coaching Process	3
Problem Statement	3
Limited Empirical Evidence of Effectiveness	4
Isolating the Effects of Executive Coaching	5
Purpose of the Study	7
Research Question	7
Significance	7
Definition of Terms	9
Delimitations	9
Limitations	10
Chapter 2: Literature Review	11
Overview of the Literature on Coaching	11
Doctoral Dissertations	11
Quantitative Empirical Studies on Executive Coaching	13
Review of Studies	13
Limitations of Studies	21
Isolating the Effects of Feedback, Training, and Coaching	23
Effects of Training and Feedback on Leadership Behavior Change	24
Effects of Feedback Alone on Leadership Behavior Change	25
Effects of Feedback and Coaching on Leadership Behavior Change	25
Chapter 3: Methods	27

Design 27
Population and Sample 29
Variables 33
Measurement Issues 34
Reliability and Validity 37
Procedures for Data Collection 37
Feedback Only (A) Group 37
Feedback and Training (B) and Feedback, Training, and Coaching (C) Groups 38
Treatment Conditions 39
Feedback Only (A) Group 39
Feedback and Training (B) Group 39
Feedback, Training, and Coaching (C) Group 40
Data Analysis 41

Chapter 4: Results 43
Demographic Data 43
Factor Analysis 45
Equivalence of Pre-test and Post-test Factor Structures 46
Identifying the Number of Factors 49
Selecting Items for Each Factor 50
Constructing Factor Scores 58
Testing the Research Question 59
MANCOVA for the Relationship Factor Scores 60
MANCOVA for the Task Factor Scores 61
Differences Among Pre-test Scores on the Task Factor 64
Differences Among Post-test Scores on the Task Factor 65
Two-Way ANOVA for Gender 66
Paired Sample *t* Tests Within Each Group 67
Summary of Results 68

Chapter 5: Discussion 71
Differences in Leadership Behavior Change across Groups 71
Differences in Relationship Factor Scores across Groups 71
Differences in Pre-test Task Factor Scores across Groups 73
Differences in Post-test Task Factor Scores across Groups 75

Decline in Post-test Scores for Feedback, Training, and Coaching (C) Group 76
Differences Between Pre-test and Post-test Scores Across Groups 78
Summary of Findings 79
Implications for Theory and Practice 80
Implications for Theory 80
Implications for Practice 86
Limitations 86
Recommendations for Future Research 91
Conclusion 92

Footnotes 93

References 94

Appendices 101
Appendix A. Demographic Data Form 102
Appendix B. Feedback Instrument 104
Appendix C. Pre-test Letter (Group A) 105
Appendix D. Sample Feedback Report 106
Appendix E. Feedback Summary (Group A1) 107
Appendix F. Post-test Letter (Groups B, C) 108
Appendix G. Scree Tests 109
Appendix H. Consent Form 115

List of Tables

		Page
Table 2.1	Effect of Feedback and Coaching on Management Competence	14
Table 2.2	Different Methods to Effect Leadership Behavior Change	24
Table 3.1	Sample Sizes and Dates of Pre-tests and Post-tests	30
Table 3.2	Variables in the Study	33
Table 3.3	Number of Questions per Subscale and Sample Questions	34
Table 4.1	Participant Age by Group	44
Table 4.2	Participant Gender and Educational and Organizational Level by Group	44
Table 4.3	Preliminary Eigenvalues for Group A Pre-test vs. Post-test Responses	47
Table 4.4	Preliminary Eigenvalues for Group B Pre-test vs. Post-test Responses	48
Table 4.5	Preliminary Eigenvalues for Groups B and C Pre-test vs. Post-test Responses	49
Table 4.6	Rotated Factor Pattern for Group A	52
Table 4.7	Rotated Factor Pattern for Group B	53
Table 4.8	Rotated Factor Pattern for Group BC	56
Table 4.9	Summary of Deleted Items	57
Table 4.10	Relationship and Task Factors	58
Table 4.11	Means and Standard Deviations of Relationship Factor Scores and Covariates	60
Table 4.12	Multiple Analysis of Covariance of Relationship Factor Scores	61
Table 4.13	Means and Standard Deviations of Task Factor Scores and Covariates	62
Table 4.14	Multiple Analysis of Covariance of Task Factor Scores	62
Table 4.15	Two-Way ANOVA for Group on Task Factor Scores	63
Table 4.16	Test of Effect Slices for Pre-test and Post-test Scores by Group	64
Table 4.17	Least Squares Means and Standard Errors of Task Pre-test Scores	64
Table 4.18	Tukey-Kramer Test of Least Squares Means for Pre-test Scores on the Task Factor	65
Table 4.19	Least Squares Means and Standard Errors of Task Post-test Scores	65

Table 4.20	Tukey-Kramer Test of Least Squares Means for Post-test Scores on the Task Factor	66
Table 4.21	Two-Way ANOVA for Gender on Relationship Difference Scores	66
Table 4.22	Two-Way ANOVA for Gender on Task Difference Scores	67
Table 4.23	Paired Sample *t* Tests by Group	68
Table 5.1	Task Factor Items	74
Table 5.2	Least Squares Means and Standard Errors of Task Pre-test and Post-test Scores	75
Table 5.3	Paired Sample *t* Tests by Groups	78
Table 5.4	*t* Tests for Participants' and Nonparticipants' Pre-test Scores in Group C	88
Table 5.5	Attrition Between Pre-test and Post-test by Group	89

List of Figures

		Page
Figure 3.1	Research Design Model	28
Figure 5.1	Center for Creative Leadership Leader Development Model	81
Figure 5.2	Three-Stage Model of Leadership Behavior Change	82

Chapter 1: Introduction

Background: Rapid Growth

Coaching has been one of the fastest-growing new trends in leadership development over the past 10 years. Most issues of professional publications in the fields of human resources and training and development, like *Training* (Bolch, 2001; Filipczak, 1998), *Training & Development Journal* (Nowack & Wimer, 1997; Thach & Heinselman, 1999; Witherspoon, White, & Hutchinson, 1996), *HR Reporter* (Peterson & Hicks, 1999), and *Human Resource Executive* (Jossi, 1999; Patton, 2001a, 2001b) include articles on coaching—defining it, exhorting people to measure its effectiveness, providing guidelines for selecting coaches, or relating coaching success stories. Even the popular press, including such magazines as *Newsweek* (Hamilton, 1996) and *Fortune* (Morris, 2000), has reported on the phenomenon.

In May 2001, *Training* magazine reported that there are some 10,000 full- and part-time coaches worldwide, according to the International Coach Federation (ICF) (Bolch, 2001). In 1996, *Newsweek* (Hamilton) reported that there were 1,000 coaches in the United States. Over a five-year period, this employment sector has grown by a factor of 10. This growth trend is not limited to the boardroom. *Training* (Bolch, 2001) reported that junior executives, newly promoted managers, and others judged to have high potential to succeed are often provided coaches. Although it used to be more common for coaches to work with managers who were in trouble or likely to "derail," coaching is now more commonly growth-oriented, helping employees develop skills and competencies that will help them succeed and advance (Bolch, 2001).

Definition of Executive Coaching

The roots of coaching are deep, and there are many types of coaching included in the field. The tradition of a sports analogy is pervasive (Kilburg, 2000). For many years, managers have been expected to coach as a vital part of their jobs (Fournies, 1987; Gilley & Boughton, 1996; Peterson & Hicks, 1996; Stowell & Starcevich, 1998). Career or outplacement coaching has been an established process provided for many employees terminated by companies during business layoffs and downsizing. The term *executive coaching* has been widely accepted as a way to differentiate a particular type of coaching. There are many definitions of executive coaching (Douglas & Morley, 2000; Hudson, 1999; Smither & Walker, 1995; Witherspoon & White, 1997). One of the most comprehensive was developed in 1996 by Kilburg, who defined coaching as a:

> … helping relationship formed between a client who has managerial authority and responsibility in an organization and a consultant who uses a wide variety of behavioral techniques and methods to help the client achieve a mutually identified set of goals to improve his or her professional performance and personal satisfaction and, consequently, to improve the effectiveness of the client's organization within a formally defined coaching agreement. (Kilburg, 1996, p. 142)

It should be noted that *executive* is frequently understood very broadly and could easily mean either manager or employee. That said, professionals in the practice of executive coaching commonly accept the following four propositions as components of the definition of coaching: (a) coaching is a one-on-one relationship between a coach and a manager or leader; (b) the coach uses a wide variety of behavioral methods in working with the client or

coachee; (c) there is a mutually agreed set of goals aimed at improving performance and satisfaction; and (d) this is bounded by a formal coaching agreement (Thach, 2002).

Purposes of Executive Coaching

In *Four Essential Ways that Coaching Can Help Executives,* Witherspoon and White (Witherspoon & White, 1997) describe four distinct purposes of executive coaching: improvement of managerial skills; correcting serious performance problems; preparing for or completing a specific challenge or task; or long-term development, often for a future leadership role in the organization.

The Coaching Process

Although there are a number of different goals, strategies, and styles used in executive coaching, the process usually has four parts: (a) goal-setting, (b) assessment, (c) awareness and action planning, and (d) implementation and monitoring (Douglas & Morley, 2000).

In the initial meeting, goals are set, and the coach forms a contract with the client. The assessment phase frequently uses 360-degree feedback, interviews, and personality measurements to gather information about the client's strengths and weaknesses. In the next phase, the coach provides the assessment information to the client in order to build the client's awareness of how others perceive him or her and assists the client in developing a personal behavior change plan. Over a period of 6 to 12 months, the client implements this action plan with interim discussions and monitoring with the coach (Douglas & Morley, 2000).

Problem Statement

Despite the phenomenal growth of executive coaching, a survey commissioned by

Personnel Decisions International reported that only 10% of companies measure the effectiveness of executive coaching (Peterson & Hicks, 1999). There is extensive anecdotal, case study, and other qualitative research on executive coaching (Crane, 2001; Dotlich & Cairo, 1999; Goldsmith, Lyons, & Freas, 2000; Hudson, 1999; Kilburg, 2000; Leonard, 1999); however, these studies describe the process of executive coaching, not its effects.

In addition, there are also a number of return-on-investment (ROI) studies (Anderson, 2001; Patton, 2001b; Peterson & Hicks, 1999) based on Phillips' (Phillips, 2003) ROI model of evaluation that report returns on investment as high as 700% for coaching. Unfortunately, these studies are based on unvalidated self-report performance data from executive coaching clients and cannot be considered as valid empirical evidence of the effectiveness of executive coaching.

Limited Empirical Evidence of Effectiveness

There are surprisingly few (Collins & Holton, 2004; Douglas & Morley, 2000; Kilburg, 2000) studies that document the effectiveness of executive coaching as measured by changed leadership behavior using accepted empirical methods. Of the few studies done, results have been conflicting. Only eight of the 10 studies reported have measured effectiveness by examining changed leadership behavior; the other two used other measures (productivity and measures of cognitive learning) to assess the effectiveness of the coaching intervention (Dore, 2001; Olivero, Bane, & Kopelman, 1997). Three of these studies did not subject their results to standard tests of statistical inference (Goldsmith, 2004; Goldsmith & Morgan, 2004; Thach, 2002).

In only one of the five remaining empirical studies that examined the effect of executive coaching on leadership behavior has executive coaching been shown to make a

consistently significant difference in changing leadership behavior (Smither, London, Flautt, Vargas, & Kuchine, 2003). Of the other four studies (Collins & Holton, 2004; Luthans & Peterson, 2003; McCauley & Hughes-James, 1994; Young & Dixon, 1996), results were mixed, with significant differences demonstrated on only two thirds or fewer of outcome measures.

In the largest study (with more than 1,200 participants), the researchers found that coaching had a significant impact on leadership behavior change, but the author himself noted that the size of the effect found was small, given the large number of participants (Smither et al., 2003). In a meta-analysis of 83 managerial leadership development programs, Collins and Holton (2004) found that objective expertise outcomes, which would include multirater feedback, were only moderately effective in studies with pre-test/post-test designs.

The practice of executive coaching continues to gain momentum, despite such scant empirical evidence of its effect on leadership behavior change. Therefore, it is essential to continue to empirically examine whether executive coaching effects positive leadership behavior change.

Isolating the Effects of Executive Coaching

Additionally, it is difficult to isolate the effect of executive coaching from other interventions (such as 360-degree feedback and training), which quite frequently accompany it (Chappelow, 2004). The executive coaching process includes assessment, which in almost all cases is carried out by administering 360-degree feedback on the leader's behavior from the leader's superior(s), peers, and subordinates (Douglas & Morley, 2000; Peltier, 2001; Smither & Reilly, 2001). Feedback alone has been reported to be an effective motivator of behavior change (Atwater, Roush, & Fischthal, 1995; Bernardin, Hagan, Ross, & Kane,

1995; Hazucha, Symanski, & Birkeland, 1992; Hegarty, 1974; J. W. Johnson & Ferstl, 1999; Reilly, Smither, & Vasilopoulos, 1996; Smither & Walker, 1995; Walker & Smither, 1999).

None of the cited studies has tested whether feedback alone leads to significantly more change in leadership behavior than feedback and coaching does. There is a large number of validated 360-degree feedback instruments available (Fleenor & Prince, 1997; Leslie & Fleenor, 1998; Tornow, London, & Associates, 1998), so feedback is an increasingly accessible and inexpensive leadership development intervention. Thus, there is a need for a controlled study of how feedback alone compares with a full executive coaching intervention in effecting change in leadership behavior.

Furthermore, most executive coaching engagements include not only these two interventions (360-degree feedback and executive coaching) but also a third intervention, attendance at a leadership development training program, which is designed to help leaders recognize effective and ineffective leadership behaviors, become positively motivated to change their leadership behavior, and set goals for changing leadership behavior. Many of these factors are designed into these training programs precisely because they have been theoretically or empirically shown to lead to behavior change (Conger, 1992; Guthrie & Kelly-Radford, 1998; Guthrie & King, 2004). Studies previously reported in the literature have not compared how feedback and training compares with a full executive coaching engagement in effecting leadership behavior change.

It is important to understand whether there are empirically validated differences between the effect of feedback and training and the effect of feedback, training, and executive coaching on leadership behavior change.

Purpose of the Study

The purpose of this study was to examine the differences in leadership behavior change due to these three different interventions: 360-degree feedback alone; 360-degree feedback and training; and 360-degree feedback, training, and executive coaching.

Research Question

The research question examined in this study was: Are there significant differences in the change scores in leadership behavior for each of the three following groups:

1. leaders who received 360-degree feedback only on their leadership behavior from their superiors, peers, and subordinates;
2. leaders who received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates, and who also attended a five-day leadership training program designed to promote leadership behavior change;
3. leaders who: (a) received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates; (b) attended a five-day leadership training program designed to promote leadership behavior change; and (c) received executive coaching, which consisted of 10 to 20 hours of one-on-one coach interaction with the client, or who received peer coaching, which consisted of 16 hours of group interaction with peers and a trained consultant?

Significance

The significance of these findings is that they will enable practitioners to better determine what combination of 360-degree feedback, training, and executive coaching is necessary to effect the desired change in leadership behavior.

If executive coaching is shown to be significantly more effective than feedback alone or feedback and training in changing leadership behavior, the effectiveness of executive coaching will have been demonstrated, and further research could investigate whether executive coaching interventions of shorter duration can be equally effective. This could lead to researchers being able to determine how many hours of coaching are necessary in order to have a significant impact on behavior change. This study had a coaching intervention of 10–16 hours, but many executive coaching interventions may involve as few as three hours (Luthans & Peterson, 2003; Smither et al., 2003; Thach, 2002). If the number of coaching hours can be reduced, the cost of executive coaching can be reduced, making it accessible to many more leaders in business and nonprofit settings.

In addition, if executive coaching is found to be significantly more effective than a combination of training and feedback, future researchers can continue to investigate whether coaching and feedback without training can be as effective as, or more effective than, the combination of the three interventions most commonly administered together (feedback, training, and coaching).

When executive coaching first began in the 1970s and ’80s, it was almost always delivered when the coach and client were face to face. In the mid-’90s, telephone coaching was introduced and rapidly replaced face-to-face meetings, after the coach and client had established a relationship in the early sessions (Leonard, 1999). Virtual coaching delivered over the Internet, using synchronous and asynchronous methods of communication, was introduced in the early 2000s (Patton, 2001a) and has been shown to be effective for some aspects of developmental relationships, such as coaching (Pulley, Sessa, & Malloy, 2002). If coaching can be found to be effective when separated from training, new delivery models can

be developed that will provide opportunities to help leaders change their leadership behavior without the need to commit time to attend a face-to-face training program. This would make the intervention more efficient and would mean that coaches could virtually accompany their clients on business trips and projects. Development would not have to be constrained to start at a fixed time and place but could be more seamlessly integrated with the professional and personal activities of leaders.

Executive coaching interventions are expensive, typically costing in excess of $20,000 per intervention (*Farr Associates*, 2003). A five-day training workshop, which includes 360-degree feedback, may cost approximately $5,000–$7,000 (*Center for Creative Leadership*, 2003; *Farr Associates,* 2003). Executive coaching interventions are lengthy, often lasting from 6 to 12 months in duration (Hall, Otazo, & Hollenbeck, 1999; Kaplan, Drath, & Kofodimos, 1991; Kaplan & Palus, 1994; Kilburg, 2000). The 360-degree feedback can be repeated at any time following the training program. If 360-degree feedback alone or 360-degree feedback with training is found as effective as executive coaching, many more leaders could benefit from these less costly and speedier interventions.

Definition of Terms

Terms, such as *executive coaching*, are defined as they are introduced in each chapter. Operational definitions are provided in chapter 3.

Delimitations

This study is limited to executive coaching as defined. Findings and recommendations cannot be applied to popular definitions of coaching, such as those employed by the ICF (*International Coach Federation Home Page*, 2004), the International Association of Coaches (ICA) (*International Association of Coaches Home Page*, 2003), or

CoachVille (*CoachVille Home Page*, 2004). The full range of extant coaching practices, as exemplified by the practices sponsored by these organizations, covers a much broader array of purposes; does not necessarily include any formal assessment of the client's behavior; seldom includes a training component; and may be of much shorter duration. These coaches often have varying levels of education, professional certification, and licenses (*CoachVille Home Page*, 2004; *International Association of Coaches Home Page*, 2003; *International Coach Federation Home Page*, 2004).

Limitations

The limited sample size used in this study may decrease the generalizability of its findings. This limitation is discussed in detail in chapter 5.

Chapter 2: Literature Review

Overview of the Literature on Coaching

Kilburg (2000) stated that there were hundreds of articles and many dozens of books on the topic of coaching in the psychological and management literature. He said the majority of this literature was devoted to athletic coaching, followed by literature focusing on the application of coaching to change problem behaviors of specific populations. Kilburg (2000) said literature focusing on coaching in a management context was the smallest body of coaching literature. Within this category, the largest area comprised books and articles that concentrated on specific coaching processes and provided anecdotal evidence of success. The second-largest area of coaching literature in the management category was focused on encouraging managers to adopt coaching techniques as a part of their management repertoire. The smallest category within the management literature on coaching was devoted to formal research on coaching (Kilburg, 2000).

Doctoral Dissertations

Research on coaching reported by Kilburg (2000) included five doctoral dissertations. Peterson (Peterson, 1993) used a customized rating inventory by participants, their bosses, and their coaches to test for changes in specific training objectives. Peterson observed 1.54 standard deviations of change on specific training objectives and only .85 standard deviations of change on a global measure of leadership effectiveness.

Kilburg (2000) reported on two other doctoral dissertations that studied of coaching as a transfer-of-training tool. One (Sawczuk, 1990) investigated a model for manager coaching of staff. The other dissertation (Miller, 1989) also studied the effect of manager coaching of participants following training. Miller found no significant differences between

the experimental and control groups after four weeks. Thompson (1986) studied participants who were at risk for career derailment and found that 40% did not complete recommended coaching and training. Those who completed all training scored significantly higher on eleven of fifteen outcome measures than those who only partially completed training. The final doctoral dissertation (Duffy, 1984) focused on the effects of coaching on outplacement. Duffy found positive but insignificant reductions in time to find a job among those who were coached.

There are 10 additional dissertations on executive coaching listed in *Dissertation Abstracts International*. Kampa-Kokesch (2001) found mixed results regarding the effect of executive coaching on leadership. Self-report data showed a positive relationship, but multiple raters showed significant differences on only part of the outcomes. Fanasheh-Husam (2003) surveyed CEOs of the top 500 American companies and found that 32% had hired coaches. Hughes (2002) developed psychotherapeutic strategies for coaching narcissistic executives. Two studies examined gender differences (Ballinger, 2000; Brown, 1999). Two qualitative studies (Orenstein, 2000; Turner, 2003) examined creating new conceptual frameworks. Three other qualitative studies (Bricklin, 2001; Charbonneau, 2002; Sztucinski, 2001) included a phenomenological perspective, an analysis of media selection, and an integration of the principles of emotional intelligence.

The Peterson (1993) and Kampa-Kokesch (2001) studies are somewhat relevant for this study in that they found some evidence of the effectiveness of executive coaching in changing leadership behavior. The relevance of Peterson's study is limited by the fact that Peterson used an index composed of coach's, boss's, and self ratings rather than multirater scores. The relevance of Kampa-Kokesch's study is limited by the facts that it had a very

small number of other raters and did not solicit feedback from participants' bosses. None of the other dissertations is relevant to this study because of methodology or variables studied.

Quantitative Empirical Studies on Executive Coaching

Review of studies. Kilburg (2000) cited only four non-dissertation research studies that investigated the effectiveness of coaching. Three of the four dealt with managers as coaches of their own employees as opposed to executive coaches, who have a different (i.e., contractual) relationship with their clients; and the outcomes examined were quite distinct from leadership behavior change, such as job-specific skills, employee retention, and career advancement.

In their annotated bibliography, Douglas and Morley (1997) used a definition of executive coaching similar to the definition used in this study. They, too, reported that they were able to identify only three published empirical studies that specifically examined the effectiveness of executive coaching in the previous 10-year period. In addition, Vicere and Fulmer (Vicere & Fulmer, 1997) reported on a compilation of research over a number of years from Keilty Goldsmith & Company. In 2002, Thach published an empirical study. In 2003, Smither et al. published an additional empirical study of the effectiveness of executive coaching. Later that year, another study (Luthans & Peterson, 2003) was published. In 2004, Collins and Horton published a meta-analytic study of managerial leadership training that included some data on coaching programs.

In addition, this researcher also identified an unpublished doctoral dissertation (Knodt, 1990) that studied the impact of coaching and a relevant current research report (Dore, 2001). Following is a review of these nine empirical studies and an analysis of their methods and findings.

In his unpublished doctoral dissertation, Knodt (1990) studied the effect of feedback and coaching on leadership behavior change. A group of 23 managers in a financial services company attended a five-day management development program based on 39 research-based behavioral people-management practices. Managers in the control group (9) were told that they would receive a self-assessment and the results of their staff members' assessment of their competence on the 39 practices as measured by a survey developed for the study. They were asked to work on improving their skills and were told that a final self-assessment and staff assessment would be administered six months later.

Managers in the experimental group (10) went through the same process but were told that staff surveys would be readministered every three weeks for six months, followed each time by one-on-one feedback and coaching sessions with the program facilitator. The improvement in competence on the 39 practices in the experimental group was significantly greater than improvement shown in the control group, as shown in Table 2.1.

Table 2.1

Effect of Feedback and Coaching on Management Competence

	Pre-test (%)	**Post-test (%)**
Experimental	16	83
Control	22	55

From *Effects of Feedback and Coaching on the Transfer of People Management Skills from the Cognitive to the Behavioral State,* by G. J. Knodt, (1990), unpublished dissertation, Columbia Pacific University, Novato, CA.

The results show that 67% of managers in the experimental group (feedback and coaching) increased their competence to the standard level of 3.75, whereas only 33% of managers in the control group (feedback only) improved their scores to standard (Knodt,

1990). No statistical tests of significance are reported; however, the results reported would be statistically significant by an ad hoc chi-square test, $p < .001$.

In 1994, McCauley and Hughes-James studied the effect of executive coaching on behavior change for 38 school superintendents following a six-day training program on leadership development at the Center for Creative Leadership. For a year following completion of the training portion of the program, superintendents worked with coaches to carry out the development plans they established as part of the training program. They met with coaches face to face and corresponded by telephone and mail at variable times during the year.

The *Benchmarks*© survey instrument was used for both pre- and post-test measurement of both self-assessment and assessment by staff (subordinates) of superintendents and board members (superordinates) chosen by superintendents. *Benchmarks*© is an internationally respected 360-degree feedback instrument developed by the Center for Creative Leadership. Scores on *Benchmarks*© have been shown to predict managerial effectiveness and career progress (McCauley, Lombardo & Usher, 1989), so it was felt to be an appropriate measure of positive program outcomes.

On pre- and post-tests one year after the conclusion of the training program, superintendents demonstrated a significant ($p < .05$) positive change score on 14 of 16 of the instruments subscales in self-assessments. Pre and post-test measures by other raters (staff or board members) chosen as observers by superintendents showed positive change on all but two subscales, of which six were significant ($p < .05$). There was no control group in this study (McCauley & Hughes-James, 1994).

Young and Dixon (1996) studied the impact of coaching on participants in another leadership development program. Participants in the *Leader Lab* program at the Center for Creative Leadership were engaged in a six-month-long leadership development program. They attended a six-day traditional classroom program followed by three months back on the job. Then they attended an additional three-and-a-half-day training program, followed by two and a half months back on the job. They worked with a coach during portions of the classroom training sessions and at least once a month by telephone during on-the-job portions.

This study of 67 middle- and upper-level managers, from a wide cross-section of business, nonprofit, and government agencies, showed significant positive ($p < .01$) change at the end of 12 months on 13 of 14 subscales of an *Impact Questionnaire* designed specifically to measure change associated with the program. The questionnaire was completed both by participants themselves and raters (co-workers) they selected. In comparison with the control group, the experimental group showed positive change on all 14 of the *Impact Questionnaire* subscales and significant ($p < .01$) change on 8 of the 14. Participants also were asked to rate which components of the program (coaching, feedback, classroom activity, etc.) were most helpful to their learning. Coaching was rated number 1 (Young & Dixon, 1996).

Prominent executive coach Marshall Goldsmith has done follow-up studies and collected data on his coaching clients over a number years and has amassed a database of more than 8,000 managers who have received some combination of feedback, training, and coaching (Goldsmith, 2004). Goldsmith's follow-up process consisted of administering a mini-survey of six questions, the first three of which were the following:

1. Do you feel this individual has become more effective (or less effective) as a leader in the past (time)?

2. Did this person discuss what he/she learned from his/her previous leadership inventory feedback with you?

3. How has this person followed up with you on areas that he/she has been trying to improve?

The first question was answered on an 11-point Likert-type scale ranging from -5 to +5. Responses to the second question were given by checking boxes designated *yes*, *no*, *not sure*, and *this person did not receive prior feedback from me*. Responses to the third question were given by checking boxes designated *consistent (periodic) follow-up*, *some follow-up*, and *no perceptible follow-up*. The other three questions in the mini-survey were selected from the original 360-degree feedback questionnaire items based on the specific behaviors the manager had decided to try to change (Vicere & Fulmer, 1997).

Goldsmith's results consistently showed that managers who met with the feedback responders after the program to thank them for their responses and who also later followed up with them to ask for feedback on their progress were more likely to be perceived as having improved in leadership effectiveness (Goldsmith, 2004; Goldsmith & Morgan, 2004; Goldsmith & Underhill, 2001; Vicere & Fulmer, 1997).

In 2002, Thach published a three-year study of 281 executives and high-potential managers in a midsized global telecommunication company in the western United States. All participants received feedback, a one-day training session, and coaching, consisting of an initial feedback session with the coach during training and three follow-up coaching sessions, as well as a mini-survey as a follow-up at six months. Her study's procedures and findings

replicated Goldsmith's (2004; Goldsmith & Morgan, 2004). Managers who did only a little follow-up were perceived to have increased in leadership effectiveness by 37%. Those who did some follow-up increased their leadership effectiveness by 57%. With frequent follow-up, perceptions of enhanced leadership effectiveness rose to 70%, and with consistent follow-up managers were perceived to have increased their effectiveness by 85% (Thach, 2002).

Smither et al. (2003) studied the effect of executive coaching and feedback on leadership behavior change in 1,361 senior managers in a large global financial services organization using a pre- and post-test 360-degree feedback design similar to that used in the two Center for Creative Leadership studies. They compared the effect of feedback and coaching with the effect of feedback alone on leadership behavior change over a 12-month period. All senior managers received 360-degree feedback in the autumn of 1999. A total of 404 senior managers worked with an executive coach after receiving their feedback. In the autumn of 2000, there was a second administration of the company-wide 360-degree feedback program. They found that managers who received feedback and coaching were significantly more likely to show improvement in feedback from direct reports and supervisors, but not from peers, as compared with managers who received feedback alone. The effect size of this difference was small (Smither et al., 2003).

Luthans and Peterson (2003) studied the effect of feedback and coaching without training on a group of 20 managers in a small manufacturing company in the Midwest. Although the focus of their study was to find out how much participants reduced the discrepancy between their self-rating and others' ratings, indirectly their data showed that managers had significant positive change on most of their behavioral scales and effected a positive change on the personal responsibility scale (Luthans & Peterson, 2003).

Collins and Holton (2004) performed a meta-analysis of 83 managerial leadership development programs, which replicated the well-known Burke and Day (1986) meta-analysis of managerial training effectiveness. They added an outcome category to those used in Burke's and Day's study, called *expertise objective* outcomes, which included "tangible results that evaluate changes in on-the-job behaviors or supervisor ratings of specific observable behaviors" (Collins & Holton, 2004, p. 225). This category included 360-degree feedback. In their analysis, they assigned studies to three different subgroups based on experimental design: post-test only with control, pre-test post-test with control, or single group retest post-test. Effect sizes ranged from a low of .35 (pre-test post-test with control) to a high of 1.01 (single group pre-test post-test); however, the authors note that only 13% of the studies were feedback interventions, and only 5% were coaching and mentoring interventions. They conclude that it is not possible to determine the effectiveness of feedback and coaching interventions because there were so few empirical studies published.

Olivero, Bane, and Kopelman (1997) studied the effects of executive coaching on productivity following management productivity training for a group of 31 managers in a United States government agency. They found that management productivity training alone accounted for increased productivity of 22.4% in managers' units. When augmented by coaching, productivity increased by 88%. The difference between 22.4% and 88% is significant at the $p < .05$ level, as measured by a *t* test (Olivero, Bane, & Kopelman, 1997).

In a comparative study of the effectiveness of different learning methodologies, Dore (2001) compared the effect on learning of executive coaching, small-group instruction, computer-based training (CBT), and traditional classroom instruction. The subjects were a group of 171 employees of a large Midwestern bank. Each subject was administered a pre-

test and post-test of the *Nelson-Denny* reading test and the *Dailey Business English* test, designed to test reading comprehension, vocabulary knowledge, grammar, punctuation, capitalization, and spelling.

Each subject received three hours of assessment prior to training sessions. Based on their needs, subjects were assigned to oral and written, interpersonal training, or all three interventions. The coach/trainer-to-trainee ratio varied across treatment groups: executive coaching (1:1), small group (1:2, 3), classroom (1:15), CBT, or control. The actual training lasted one hour per week for either 10 or 20 weeks.

The executive-coaching treatment produced five dependent variable scores significantly ($p < .01$) higher than the small-group and classroom methodologies. One dependent variable score for the small-group methodology was higher than the classroom methodology. No significant differences were found for the CBT or classroom treatment groups (Dore, 2001).

The Dore and Olivero et al. studies are perhaps less important to this review than the previous eight because they focus on different dependent variables, learning, and productivity increases. The Olivero study in a public agency included in its training intervention a combination of behavioral skills (i.e., management skills) and cognitive knowledge (i.e., productivity improvement processes through measurement of inputs and outputs) (Olivero, Bane & Kopelman, 1997). One confounding variable that may be more troubling in the Olivero et al. study is the requirement that participants provide a public oral report of project results to their managers. Such a design factor in the intervention doubtless adds accountability as a strong incentive for the learner; however, it may obscure the effects

of coaching and could be an impediment to experimenting with new behavior, which may be important for making changes in more complex behavioral skill sets.

The Dore study of banking employees included primarily cognitive knowledge outcomes (reading comprehension, vocabulary knowledge, grammar, punctuation, capitalization, and spelling) as its dependent variable (2001). While interesting, the Olivero et al. and Dore studies are not comparable to the other research studies that focus on leadership behavior. Additionally, neither of them used 360-degree feedback as part of their intervention. For these reasons, they were eliminated from further analysis.

Limitations of studies. All studies cited showed positive effects of coaching, many of which were statistically significant. The studies did reflect a number of different models, research designs, and variables. Following is an analysis of some of the issues and constraints these differences presented.

The study of school superintendents did not include a control group (McCauley & Hughes-James, 1994), which makes it difficult to generalize these findings. However, the 1996 study by Young and Dixon found a very similar pattern of results and was based on a very similar intervention strategy. Both were feedback-intensive programs conducted by the Center for Creative Leadership, which included 360-degree feedback, training, and coaching following the training. Based on these facts, it would seem reasonable to conclude that these findings are valid.

The Goldsmith (2004; Goldsmith & Morgan, 2004) and Thach (2002) studies utilized the same mini-survey format. In none of these studies were results analyzed using standard statistical tests of inference, although the numerical results reported did appear to be highly significant. While this assessment of increase in leadership effectiveness was based on the

perceptions of a number of individuals and not the participant's own assessment, it was based on a single global measure rather than a behavioral scale composed on a number of items. Such global measures are suspect in empirical research. In addition, no published accounts of these studies included any analysis of the respondents' scores on the three items selected from the original 360-degree survey questionnaire (the items concerning areas in which participants chose to try to change behavior). If available, these data could validate the findings on the global outcome measure. Such an analysis would enhance the credibility of these findings.

Many of these studies were based on small sample sizes: 20 (Luthans & Peterson, 2003), 23 (Knodt, 1990), 38 (McCauley & Hughes-James, 1994), and 67 (Young & Dixon, 1996). Both of the Center for Creative Leadership studies found statistically significant effects for only half of the subscales on their feedback instruments. Luthans' and Peterson's (2003) results were also mixed. The Smither et al. study (2003) found statistically significant effects for managers who received feedback and coaching compared to managers who received feedback alone as measured by direct report and supervisor feedback but not in feedback from peers. In addition, the authors themselves report that using Cohen's d statistic, the effect size of these differences was small.

Unlike all other feedback-based studies, in the study by Smither et al. (2003), feedback reports were shared with participants' supervisors, which the authors pointed out may have enhanced the perceived accountability participants felt for changing their behavior. This may make it difficult to directly compare the results of this study with others in which feedback was kept confidential and shared only with participants themselves and their coaches. Small sample sizes, the failure to achieve consistent statistical significance on all

measures, the use of global measures, and neglecting to report standard tests of statistical inference make it difficult to generalize the results of these studies on the effectiveness of coaching on changing leadership behavior with confidence.

Isolating the Effects of Feedback, Training, and Coaching

As noted earlier, executive coaching is inextricably intertwined with 360-degree feedback and almost as closely tied to leadership development training programs (Chappelow, 2004; Guthrie & Kelly-Radford, 1998; Guthrie & King, 2004). It is difficult to isolate the effect of feedback from that of training and/or executive coaching (Chappelow, 2004). These studies include some with feedback alone; some with feedback and coaching; some with feedback and training; and some with feedback, training and coaching. Table 2.2 shows the different methods employed in these studies.

Table 2.2

Different Methods to Effect Leadership Behavior Change

Study	**Feedback alone**	**Feedback + coaching**	**Feedback + training**	**Feedback + training+ coaching**
Knodt (1990)			√	√
McCauley & Hughes-James (1994)				√
Young & Dixon (1996)				√
Goldsmith et al. (2004)		√	√	√
Thach (2002)				√
Smither et al. (2003)	√	√		
Luthans & Peterson (2003)		√		

Effects of Training and Feedback on Leadership Behavior Change. The value of 360-degree feedback combined with classroom training in motivating behavioral change in management and leadership development has been firmly established by a number of researchers (Conger, 1992; Guthrie & Kelly-Radford, 1998; Guthrie & King, 2004; Vicere & Fulmer, 1997). Effective feedback-intensive programs are carefully designed to maximize the amount of assessment, challenge, and support given to the client in order to maximize the amount of behavior change (Guthrie & King, 2004). The Knodt (1990) study of financial services managers and the two Center for Creative Leadership studies (McCauley & Hughes-

James, 1994; Young & Dixon, 1996) included interventions that included 360-degree feedback within a feedback-intensive program, followed by coaching. Only one study (Knodt, 1990) compared the effect of feedback and training with the effect of feedback, training, and coaching; and this study is fraught with other problems, such as small sample size and lack of statistical tests of results.

Effects of Feedback Alone on Leadership Behavior Change. The value of feedback alone in changing leadership behavior has been documented in the literature (Atwater et al., 1995; Bernardin et al., 1995; Hazucha et al., 1992; Hegarty, 1974; J. W. Johnson & Ferstl, 1999; Reilly et al., 1996; Smither & Walker, 1995; Walker & Smither, 1999). In addition to being used in development contexts like most of the studies cited here, 360-degree feedback has been used extensively in performance evaluation settings, and a wealth of research attesting to the effectiveness of 360-degree feedback in changing managers' self-perceptions and their leadership behavior has been amassed (Bracken, Timmreck, & Church, 2001; Edwards & Ewen, 1996; London, 2001; London & Tornow, 1998). Smither and Reilly (2001) analyzed 13 research studies on the effect of 360-degree feedback on improvement in subsequent feedback scores. They found that 11 of 13 studies found evidence of improvement in leadership behavior as measured by 360-degree feedback. The Smither et al. (2003) study examined the effect of feedback and coaching and the effect of feedback alone on leadership behavior change, but they did not compare these effects with that of feedback, training, and coaching on the same dependent variable.

Effects of Feedback Alone and Coaching. Only one study (Smither et al., 2003) compared the effect of feedback alone with the effect of feedback and coaching. The Goldsmith (2004) studies were aggregated across all three conditions, so there is no way to

compare results for each condition. These studies did not contain feedback-only or control groups, either.

Chapter 3: Methods

Design

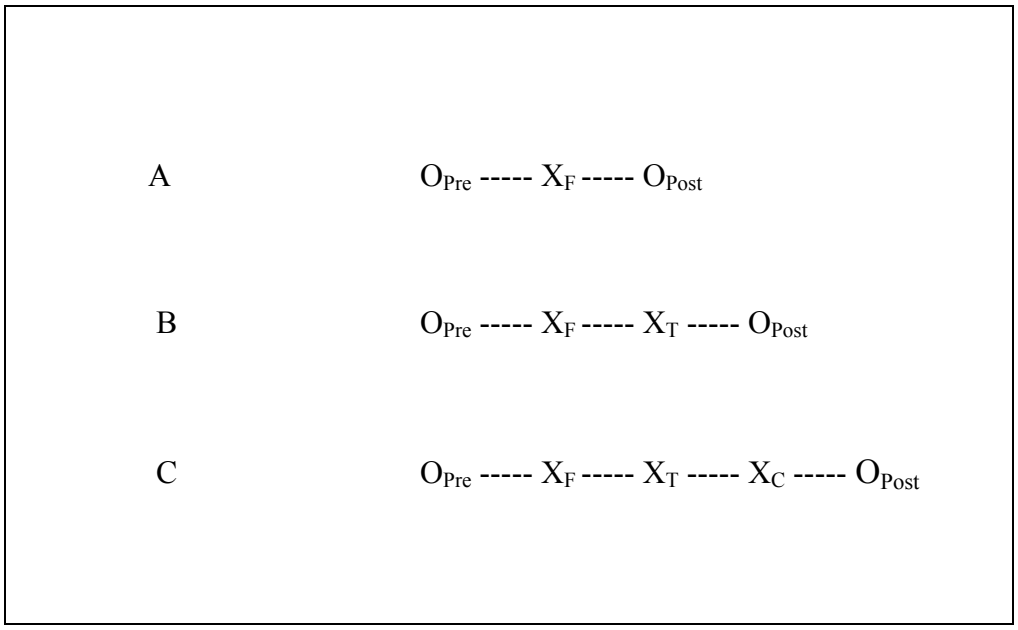
A quasi-experimental pre-test/post-test design was used for this study to investigate the research question: Are there significant differences in the change scores in leadership behavior for each of the three following groups:

1. leaders who only received 360-degree feedback on their leadership behavior from their superiors, peers, and subordinates;
2. leaders who received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates, and who also attended a five-day leadership training program designed to promote leadership behavior change;
3. leaders who: (a) received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates; (b) attended a five-day leadership training program designed to promote leadership behavior change; and (c) received executive coaching (which consisted of 10 to 20 hours of one-on-one coach interaction with the client) or peer coaching (which consisted of 16 hours of group interaction with peers and a trained consultant).

The three groups in the study are referred to as feedback only (A), feedback and training (B), and feedback, training, and coaching (C).

The dependent variable was change in leadership behavior. Operational definitions and treatment conditions are described in later sections of this chapter.

Figure 3.1. Research design model.



The research design, depicted in Figure 1, shows that in the group that received feedback only (A), participants distributed surveys concerning their leadership behavior to a group of respondents composed of their superior(s), peers, and subordinates (O_{PRE}). They received the aggregated results of the survey in a written, one-page feedback report (X_F). For the post-test measure (O_{POST}), the leadership behavior surveys were redistributed to the same or a similar group of colleagues[1] six to nine months after the administration of the pre-test. Post-test results were also distributed to participants in the study as a courtesy; however, this was outside the parameters of this study.

In the group receiving feedback and training (B), participants distributed surveys concerning their leadership behavior to a group composed of their superior(s), peers, and subordinates (O_{PRE}). They received the aggregated results of the survey in a written, one-page feedback report (X_F). They also attended a five-day leadership development training

program (X_T). For the post-test measure (O_{POST}), the leadership behavior surveys were redistributed to the same or a similar group[1] of colleagues from eight to 38 months after the administration of the pre-test. Post-test results were also distributed to participants as a courtesy; however, this was outside the parameters of this study.

In the group that received feedback, training, and coaching (C), participants distributed surveys of their leadership behavior to a group composed of their superior(s), peers, and subordinates as the pre-test measure (O_{PRE}). They received the aggregated results of the survey in a written, one-page feedback report (X_F). They attended a five-day leadership development training program (X_T) and received 10 to 20 hours of individualized executive coaching or 16 hours of peer coaching (X_c). For the post-test measure (O_{POST}), the leadership behavior surveys were redistributed to the same or a similar group[1] of colleagues from four to 32 months after the administration of the pre-test. Post-test results were also distributed to participants as a courtesy; however, this was outside the parameters of this study.

Population and Sample

The population was 2,837 clients of a small leadership development consulting firm in the southeastern United States specializing in leadership development and executive coaching from 1998 to 2004. Clients were typically middle or senior managers of medium or large for-profit businesses; clients ranged in age from 31 to 55. Participants in the feedback only group (A) were recruited from client companies and similar organizations in the Southeast United States.

The final study had a convenience sample of 101 eligible participants. There were 71 in the sample for the feedback only (A) group; 21 in the feedback and training (B) group; and 9 in the feedback, training, and coaching (C) group.

The researcher had the primary responsibility for enrolling participants in the A group. Human resources representatives from 14 organizations were contacted. The organizations were previous client companies of the sponsoring company or similar size organizations in the region. Human resources representatives contacted varying numbers of employees to invite them to participate in the study. The number of people who received this initial invitation was not tracked in this study. These participants were offered the opportunity to receive feedback twice, but no training or executive coaching, over a six- to nine-month period. One hundred and sixty-four participants agreed to participate in the pre-test (first round of feedback) administered from August to December 2004. A total of 71 of these participants also completed the post-test (second round of feedback) administered from March to July 2005. Participants received feedback reports from 5.5 to 9 months apart. Table 3.1 shows the initial numbers of participants who completed both pre-tests and post-tests and the dates the pre-tests and post-tests were administered.

Table 3.1

Sample Sizes and Dates of Pre-tests and Post-tests

Group	**Pre-test**		**Post-test**	
	Number	Dates	Number	Dates
A	164	Aug.–Dec. 2004	71	Mar.–Jul. 2005
B	70	Aug. 1998–Jun. 2004	36	Aug. 2004–Jan. 2005
C	20	Mar. 1999–May 2005	12	Aug. 2004–Sep. 2005

The 21 participants in the feedback and training (B) group were a convenience sample selected from more than 2,000 members of the sponsoring company's client population who completed the pre-test and received training from August 1998 through June 2004. Randomly selected individuals were contacted by a marketing representative of the sponsoring organization from June 2004 to January 2005 and offered the opportunity to receive feedback from up to 15 respondents of their choice. The post-test (second round of feedback) was administered from nine to 38 months following the pre-test and attendance at the five-day training program. A total of 70 clients agreed to take part in the study; however, only 36 completed the process.

The nine participants in the executive coaching group (C) were a convenience sample of executive and peer coaching clients of the sponsoring organization. The initial group was recruited from the pool of 31 executive coaching clients of the sponsoring organization from March 1999 to May 2005 (pre-test dates). All of these clients received feedback, training, and coaching. Consultants who provided the training and coaching services to these clients contacted them from June 2004 through May 2005 and offered them the opportunity to receive feedback from up to 15 respondents of their choice. The post-test (second round of feedback) was administered from August 2004 through September 2005, which was four to 32 months following their pre-tests, attendance at the five-day training program, and individual 10-20 hour coaching session(s). A total of 13 clients agreed to participate in the study. Six of these completed the process.

In order to recruit additional participants in the feedback, training, and coaching group (C), clients who participated in two peer-coaching follow-up programs in February and May 2005 were also offered the opportunity to participate in the study. Instead of individual

coaching, these participants attended a two-day group feedback session conducted by one of the lead organization consultants. Of the nine who were qualified for this study, seven agreed to participate, and six completed the post-test. This subgroup had taken pre-tests from December 2002 to June 2004. Post-tests were administered to this group from June through August 2005.

Of the 36 participants in the feedback and training (B) group, 15 were discarded from the study because they completed their pre-tests prior to February of 2001. Individual pre-test item scores were not available prior to that date because of a reporting and data input revision made at that time. Three participants were discarded from the executive coaching clients in the feedback, training, and coaching (C) group because pre-test results were from prior to February 2001. As a result, all eligible participants in the study completed their pre-tests and training after February 2001. Attrition and other issues affecting the small size of the number of participants in this group will be discussed in Chapter 5. The small number of eligible participants in this group may represent a limitation on generalizability of this study.

Age of participants and the length of time between administration of pre-tests and post-tests were controlled statistically through a multiple analysis of covariance (Rawlings, Pantula, & Dickey, 1998). Consultants at the organization reported that executive coaching clients (C) tended to be older and more often male than clients who received feedback and training (B). However, these data were not routinely collected as a part of the sponsoring company's normal operating procedures. The amount of time between pre-test and post-test ranged from four to 38 months. Scholars differ on the amount of time it takes to make behavioral changes following interventions like feedback, training, and coaching (Chappelow, 2004; Dalton, 1997); however, such a large range could seriously confound

results if not controlled for statistically. Age, sex, and other demographic data were collected in the post-test administration. The demographic data form is shown in Appendix A.

Variables

Table 3.2 describes the independent and dependent variables in the study, as well as the covariates. It also describes the type of data each variable represents.

Table 3.2

Variables in the Study

Variable Name	**Type of Variable**	**Type of Data**
Independent	Feedback only (A)	Nominal
	Feedback and training (B)	Nominal
	Feedback, training and coaching (C)	Nominal
Dependent	Leadership change score	Continuous
Covariates	Age	Continuous
	Time	Continuous

The sponsoring company's proprietary feedback instrument, shown in Appendix B, was used to collect pre-test and post-test leadership behavior data on participants. This instrument is a 28-item Likert-type scale composed of nine subscales, shown in Table 3.2.

Each question was measured on a range from zero (lowest) to 10 (highest). Each subscale measured commonly accepted leadership constructs, including the sample items shown in Table 3.3. Individual items are descriptions of behaviors typical of each construct.

Table 3.3

Number of Questions per Subscale and Sample Questions

Subscale title	Number	Sample question
Communication	2	Demonstrates effective listening and questioning skills.
Interpersonal Skills	4	Earns trust and loyalty.
Teamwork	5	Places the team before personal interest.
Initiative	3	Seeks new challenges and responsibilities.
Creativity	1	Develops innovative ideas and approaches that work.
Adaptability	3	Flexible in dealing with different viewpoints and styles.
Judgment	5	Considers alternatives and consequences before making a judgment.
Leadership	3	Provides purpose, values, vision.
Maturity	2	Effectively deals with stress and frustration.

The leadership change score was operationally defined as the difference between pre-test and post-test average scores of selected items from the feedback instrument. Change in leadership behavior was the dependent variable in this study. It was calculated by subtracting the average pre-test score on selected items from the average post-test score on selected items. Items were selected based on factor loading obtained from exploratory factor analyses.

Measurement Issues

There are four measurement issues associated with multisource feedback: response shift bias, regression to the mean, ceiling effect, and difference scores (Smither & Walker, 2001). Response shift bias is the tendency for responses to shift, often downward, upon

repeated measurements using the same scale. In essence, the theory is that as attention is focused on the criteria raters' judgments, their judgments become more refined and more critical (Martineau, 1998). Several approaches to dealing with response-shift bias have been suggested. One approach is to use retrospective pre-tests. For example, raters are asked to make two ratings at time 2. The first rating asks the rater to evaluate retrospectively the effectiveness of the person's behavior at time 1. The second rating asks the rater to evaluate the effectiveness of the person's behavior now (at time 2) (Smither & Walker, 2001).

Another method for dealing with response-shift bias involves using a retrospective degree-of-change rating. In this approach, change is assessed directly rather than being inferred from a comparison of pre-test and post-test (time 1 versus time 2) scores. For example, at time 2, raters can be asked to rate the person's current degree of change. Martineau (1998) describes several studies that found that retrospective pre-tests or retrospective degree-of-change ratings yielded more accurate representations of change than did traditional pre-tests (such as comparing time 1 and time 2 ratings). Other researchers contend that response-shift bias applies only to self-ratings and not to ratings by others (Craig, 2002) and have demonstrated that item-response theory can be accurately used for multisource behavioral data, despite the violation of the assumption of independence (Craig & Kaiser, 2003).

Regression to the mean is the tendency for scores on unreliable instruments to approach the mean upon a second administration of the instrument. Many feedback instruments have questionable reliability. One consequence of regression to the mean is that those who score very high on the first administration of a somewhat unreliable measure are likely to score somewhat lower on a second administration of the measure. Conversely, those

who score very low on the first administration of a somewhat unreliable measure are likely to score somewhat higher on a second administration of the measure (Smither & Walker, 2001). Regression to the mean can be controlled statistically by using the reliability of the instrument to estimate how much scores would change solely due to regression to the mean; then, only scores in excess of this measure would be used for analysis (London & Smither, 1995). However, the most recent research on the reliability of multisource instruments suggests that most of the variance in ratings can be explained by rater idiosyncrasy (Craig & Kaiser, 2003; Yammarino, 2003), which casts serious doubt on the usefulness of reliability based on previous research studies. Therefore, controlling for regression to the mean using this technique was not considered for this study.

Ceiling effect is the phenomenon that occurs when all ratings are clustered in the high region of the scale. If all ratings average 4.8 on a 5-point scale, there is little room below the ceiling in which to demonstrate an improvement, so the actual change may be masked by the ceiling effect (London & Smither, 1995). Most rating scales for 360-degree feedback range from four to seven points. It is unclear whether increasing scales to 10 points would increase the accuracy of the ratings (Smither & Walker, 2001). Since this study used a pre-existing multisource instrument that was based on a 10-point scale, ceiling effect was minimized.

The concept of difference scores suggests that the difference between self-ratings and other ratings influences behavior change. For example, a manager with a large difference between the two would be more motivated to change, according to this theory (Smither & Walker, 2001). Polynomial regression is the method of choice for addressing this problem (J. W. Johnson & Ferstl, 1999). The feedback collection and reporting procedure used by this

organization did not include self-ratings; therefore, polynomial regression was not considered.

Reliability and Validity

Validity and reliability analyses had not been done on the feedback instrument prior to this study. Recent research has shown that the actual factors identified in multisource leadership behavior data rarely match the factors identified in earlier data collections using the same instrument (Craig & Kaiser, 2003). For this reason, exploratory factor analyses (Hatcher & Stepanski, 1994a) were performed on the pre-test and post-test responses, and only items that demonstrated high primary-factor loading (.40 or above) and low cross-factor loading (loading on multiple factors) were selected for data analysis. The procedure described in Craig and Kaiser's (2003) study of the application of item response theory to multisource ratings was used. These data are reported in Chapter 4.

Procedures for Data Collection

Feedback only (A) group. Pre-test leadership behavior data for the feedback only (A) group was collected from up to 15 subordinates, peers, and superiors of each participant from August to December 2004. A copy of this letter is included in Appendix C. Feedback was reported to participants in a one-page feedback report, which included a chart that graphically showed the item means clustered by subscales. A copy of the feedback report is included in Appendix D. Subscale scores were also summed and averaged and reported to participants as means from 0 to 10. Post-test leadership behavior data were obtained for the feedback only (A) group from March to July 2005. Participants received post-test feedback as a courtesy; however, this was outside the parameters of this study.

Feedback and training (B) and feedback, training, and coaching (C) groups. For the feedback and training (B) group and the feedback, training, and coaching (C) group, pre-test leadership behavior data was collected in advance of training sessions that participants attended between February 2001 and June 2004 using the same procedures as those used for the feedback only (A) group. A letter asking each participant to distribute copies of the feedback instrument to respondents was sent to each participant two to six weeks before the training event. Names of respondents to whom participants distributed leadership behavior surveys were not reported back to the sponsoring consulting organization. The organization tracked the amount of feedback received and contacted clients to ask them to follow up with respondents to ensure adequate numbers of respondents. Pre-test leadership behavior scores were summed and averaged by item and reported to participants using the same one-page feedback report that was used for the feedback only (A) group. This is shown as Appendix D.

Post-test leadership behavior ratings were obtained using the same procedures used for the pre-test in the feedback only (A) group. Participants were informed of the opportunity to participate in a special follow-up study. Post-test data were collected for the feedback and training (B) group and the feedback, training, and coaching (C) group from August 2004 to September 2005, timed to ensure a minimum of six months and a maximum of approximately 36 months after administration of the pre-test. Participants received post-test feedback as a courtesy; however, this was outside the parameters of this study. Demographic questions, including age and sex, were also included in the post-test data administration.

Treatment Conditions

Feedback only (A) group. From August to December 2004, participants in this group received a feedback report on their leadership behavior with a letter from the sponsoring consulting firm thanking them for participating in the research study. The letter stated that feedback is helpful in developing and enhancing leadership skills and provided tips for interpreting. A copy of the letter is reproduced in Appendix E. From February to May 2005, participants in this group received another letter (shown in Appendix F) asking them to distribute post-test leadership surveys to the same individuals who responded to the pre-test.

Feedback and training (B) group. Participants received feedback on the same instrument using the same reporting procedure as used for the feedback only (A) group. They received feedback reports on their pre-test leadership behavior from February 2001 through June 2004 as part of the five-day training seminar they attended. The seminar devoted a full day to discussions of the importance of feedback. Participants met individually with an experienced consultant to assist them in interpreting their feedback data. They also discussed their feedback and their reaction to it with other participants attending the seminar. As part of the seminar, participants developed action plans for desired changes in their leadership behavior. Effective leadership behaviors were presented, discussed, modeled, and practiced during the seminar. Participants were briefed on how to discuss their feedback with their superior(s) and colleagues and encouraged to share their data and action plans with their colleagues. The training sessions were facilitated by two of the four senior consultants of the sponsoring consulting firm, who conducted approximately 40 sessions per year. Participants received a letter (shown in Appendix F) to request that they distribute post-test feedback

surveys from August 2004 to August 2005, eight to 38 months following the collection of their pre-test feedback.

Feedback, training, and coaching (C) group. The participants in this group received feedback on the same instrument using the same reporting procedure. As a part of the five-day training seminar they attended, they received feedback reports on their pre-test leadership behavior from February 2001 to May 2005. They received post-test feedback results in August 2004 to September 2005, four to 32 months following pre-test feedback. In addition, members of the original executive coaching subgroup received coaching consisting of: data from personal interviews with individuals who responded to the feedback instrument performed by an experienced executive coach; presentation by the coach of verbatim reports concerning feedback and performance issues; consultation on interpretation of impact of feedback on relationships and performance of participants and their team members; assistance with goal-setting for behavior change and action planning for behavioral change process; and interim/progress reports at 6 to 12 months. The executive coaching consisted of from 10 to 20 hours of one-on-one interaction between the participant and his or her executive coach. This definition of executive coaching is consistent with Kaplan's and Palus's definition of "enhanced feedback" (1994). Executive coaching was delivered within one year of participants' attendance at the training program. Four senior consultants from the consulting firm served as executive coaches.

Members of the peer coaching subgroup received feedback reports on their pre-test leadership behavior from December 2002 to November 2004 as a part of the five-day training seminar they attended. They received post-test feedback results from June to September 2005, nine to 32 months following pre-test feedback. In the peer coaching sub-group,

participants did not receive reports of verbatim interviews of colleagues, but they participated in a two-day follow-up group meeting with other participants who had attended the original feedback and training intervention. The group meeting included peer observations and peer feedback, and it gave attendees an opportunity to share information about strengths, weaknesses, and progress on development plans, as well as to update and strategize new development plans (Baumgartner).

Data Analysis

The research question examined in this study was: Are there significant differences in the change scores in leadership behavior for each of the three following groups:

1. leaders who received 360-degree feedback only on their leadership behavior from their superiors, peers, and subordinates;
2. leaders who received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates, and who also attended a five-day leadership training program designed to promote leadership behavior change;
3. leaders who: (a) received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates; (b) attended a five-day leadership training program designed to promote leadership behavior change; and (c) received executive coaching, which consisted of 10 to 20 hours of one-on-one coach interaction with the client, or who received peer coaching, which consisted of 16 hours of group interaction with peers and a trained consultant?

Multivariate Analysis of Covariance (MANCOVA) was used to test the differences in leadership behavior change over time and to control for age and time between pre-test and

post-test administration for the three quasi-experimental groups (Littell, Fruend, & Spector, 1991; Rawlings et al., 1998).

Chapter 4: Results

In this chapter, the results of statistical analysis to answer the research question are presented. All data analysis for this paper was generated using SAS software, Version 8 of the SAS System for Unix, copyright © 1999-2001 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA. Following the descriptive statistics of the groups of study participants are the factor analyses, the multivariate analyses of covariance, and some additional analyses that help answer questions that were raised in the initial analysis. The factor analysis resulted in the identification of two factors. Only items that loaded cleanly on either of these two factors were chosen to construct factor scores (Hatcher & Stepanski, 1994a) that were used to test the research question about differences between the groups using the multiple analyses of covariance method (Littell et al., 1991; Rawlings et al., 1998). Differences in average difference scores for each factor were analyzed, and additional effect tests and tests of differences of least squares means (Littell et al., 1991) procedures were used to examine the extent to which pre- and post-test differences among the groups may account for significant differences found. Finally, *t* tests for paired samples (Agresti & Finlay, 1997) were calculated to determine differences between pre-test and post-test scores within each group.

Demographic Data

Prior to this study, the sponsoring consulting firm had not collected any descriptive demographic information on participants in training and coaching services. In this study, selected demographic data were collected from participants in order to determine if differences in the composition of the groups affected results. Age was a covariate in this

study. The average age of all 101 study participants was 40.5, with a minimum of 26 and a maximum of 57. Consultants at the firm had estimated that clients ranged in age from 31 to 55. These data are shown in Table 4.1.

Table 4.1

Participant Age by Group

Group	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
A	71	40.5633803	7.6582960	26.0000000	57.0000000
B	21	39.5714286	7.1033191	31.0000000	55.0000000
C	9	42.1111111	5.8404718	33.0000000	50.0000000
Total	101				

On average, the feedback and training (B) group was the youngest at 39 and a half years old. The feedback only (A) group was a year older and the feedback, training and coaching group was the oldest at a little over 42 years old. On some other demographic variables the groups differed more markedly. These data are shown in Table 4.2

Table 4.2

Participant Gender and Educational and Organizational Levels by Group

Group	N	Gender (%)		Education (%)				Organization Level (%)			
		M	F	HS	HS+	Col	Grad	Top 2	Upper	Mid	Non
A	71	36	64	9	23	46	22	6	18	74	1
B	21	64	36	5	9	45	41	14	45	32	9
C	9	67	33	-	-	13	87	75	13	12	-
Total	101										

In terms of gender, the feedback only (A) group was the mirror opposite of the feedback and training (B) group. The feedback only group was predominantly female (64%); the feedback and training (B) and feedback, training, and coaching (C) groups were only about one-third female. The feedback only (A) group was predominantly (46%) a college graduate (Col) population, with 23% and 22%, respectively, having some college (HS+) and some graduate (Grad) study beyond the baccalaureate degree. The feedback and training group (B) differed markedly on the level of education. More than 80% had four-year degrees or graduate study. The feedback, training, and coaching group (C) continued the trend toward higher levels of education, with 87% having graduate study or degrees beyond the baccalaureate.

Levels in the organization were also different across the groups. The feedback only (A) group was primarily (74%) a middle management group (Mid), defined as managers, supervisors, professional staff, and team leaders. The feedback and training (B) group had twice as many (14%) in the top two organizational levels, defined as chief level officers or presidents, senior vice presidents, executive presidents, and directors. They also had almost three times (45%) as many individuals in the upper-middle tier, defined as department heads and plant managers. Almost one third (32%) of this group considered themselves middle managers, compared with almost three quarters (74%) of the feedback only (A) group. The feedback, training, and coaching group again continued this trend toward higher levels in the organization, with 75% in the top two tiers.

Factor Analysis

The consulting firm's proprietary feedback instrument, shown in Appendix B, was used to collect pre-test and post-test leadership behavior data on participants. This instrument

was an unvalidated 28-item Likert-type scale composed of nine subscales. Each item was measured on a range from zero (lowest) to 10 (highest). As is common practice in survey scale construction, each subscale was designed to measure commonly accepted leadership constructs. Individual items were descriptions of behaviors typical of each construct.

Recent research has shown that the actual factors identified in multisource leadership behavior data rarely match the factors used to develop the instrument or identified in earlier data collections using the same instrument (Craig & Kaiser, 2003). For this reason, exploratory factor analyses (Hatcher & Stepanski, 1994a) were performed on the pre-test and post-test data, with the following objectives in mind: (a) to establish that factors remained stable across pre-test and post-test administrations, (b) to identify how many actual factors accounted for significant variance in the data, and (c) to identify the items that loaded on the actual factors consistently. Subscales were constructed for each factor, and items that did not load consistently on the factors were discarded from further analysis.

Equivalence of Pre-test and Post-test Factor Structures. Responses to the 28-item survey were segregated by group (A, B, or C) and by pre-test and post-test responses. Responses were then subjected to separate exploratory factor analyses using squared multiple correlations as prior communality estimates. The maximum likelihood estimation method was used to extract the factors (Hatcher & Stepanski, 1994a). Only responses for participants who took part in both the pre-test and post-test were included in the analysis.

For the feedback only (A) group, there were 764 usable responses in the pre-test pool and 647 in the post-test. Both pre-test and post-test data for the feedback only (A) group rotated to 6 factors. The preliminary eigenvalues for the pre-test and post-test responses are compared in Table 4.3. The preliminary eigenvalues remained much the same across the pre-

and post-test administrations. The first factor was within one point of cumulative value. The second factor was within two points. In the absence of a standardized statistical test for differences, an examination of the eigenvalues of the factors showed that these factors remained stable across pre-test and post-test administration.

Table 4.3

Preliminary Eigenvalues for Group A Pre-test vs. Post-test Responses

	Pre-test				**Post-test**		
Factor	Eigenvalue	Proportion	Cumulative	Factor	Eigenvalue	Proportion	Cumulative
1	79.69	.87	.87	1	90.73	.88	.88
2	6.55	.07	.94	2	5.38	.05	.94
3	2.33	.03	.96	3	3.00	.03	.96
4	1.89	.02	.98	4	1.78	.02	.98
5	1.46	.02	.99	5	1.56	.02	.99
6	1.09	.01	1.00	6	1.23	.01	1.00

For the feedback and training (B) group, there were 215 usable responses in the pre-test pool and 219 in the post-test. Both pre-test and post-test data for the feedback and training (B) group rotated to 7 factors. The preliminary eigenvalues for the pre-test and post-test responses are compared in Table 4.4. The first factor had a similarly high value in the pre-test, but dropped in value in the post-test administration from .82 to .70. The second factor accounted for more of the variance in the post-test (.13) compared to the pre-test (.09). In the absence of a standardized statistical test for differences, an examination of the eigenvalues of the factors showed that the first few factors accounted for such a large amount of the proportion of the variance (well over 80%) that use of multiple factors resulted in stable measurement across pre-test and post-test administration.

Table 4.4

Preliminary Eigenvalues for Group B Pre-test vs. Post-test Responses

	Pre-test				Post-test		
Factor	Eigenvalue	Proportion	Cumulative	Factor	Eigenvalue	Proportion	Cumulative
1	90.34	0.82	0.82	1	52.51	.70	0.70
2	9.31	0.09	0.90	2	9.81	.13	0.83
3	3.76	0.03	0.94	3	5.86	.08	0.91
4	2.44	0.02	0.96	4	2.55	.04	0.94
5	1.85	0.02	0.98	5	1.84	.04	0.96
6	1.62	0.02	0.99	6	1.46	.02	0.98
7	1.41	0.012	1.00	7	1.37	.01	1.00

For the feedback, training, and coaching (C) group, there were 91 usable responses in the pre-test pool and 101 in the post-test. This number was inadequate for a valid exploratory factor analysis, which requires five times the number of items in the survey (28 X 5 = 140) (Hatcher & Stepanski, 1994a); therefore, groups B and C were combined. There was a total of 306 usable responses in the pre-test and 320 in the post-test for the combined B and C groups. Both pre-test and post-test data for the feedback, training, and coaching (C) group rotated to 7 factors. The preliminary eigenvalues for the pre-test and post-test responses are compared in Table 4.5. Because of the contribution of Group B in this pool of responses, a pattern similar to that shown in Group B was evident, although it was not as pronounced. The first two factors accounted for a large proportion of the cumulative eigenvalues (.91 and .83 respectively). In the absence of a standardized statistical test for differences, an examination

of the eigenvalues of the factors showed that factor structure was sufficiently stable across pre-test and post-test administrations.

Table 4.5

Preliminary Eigenvalues for Groups B and C Pre-test vs. Post-test Responses

	Pre-test				**Post-test**		
Factor	Eigenvalue	Proportion	Cumulative	Factor	Eigenvalue	Proportion	Cumulative
1	73.10	0.82	0.82	1	62.02	0.77	0.77
2	8.96	0.10	0.92	2	8.90	0.11	0.88
3	2.70	0.03	0.95	3	4.79	0.06	0.94
4	2.00	0.02	0.97	4	2.36	0.03	0.96
5	1.67	0.02	0.98	5	1.47	0.02	0.98
6	1.18	0.01	0.99	6	1.32	0.02	0.99
7	1.03	0.01	1.00	7	1.04	0.01	1.00

In order to focus on alpha change, gamma and beta change were minimized by limiting this analysis to only those items that functioned equivalently in both pre-test and post-test. Testing the research question using the difference in pre-test and post-test scores over time was, therefore, appropriate.

Identifying the Number of Factors. After the factors were extracted, a promax (oblique) rotation was used to rotate to two factors (Hatcher & Stepanski, 1994a). Separate scree tests for each group (A, B, and C) were performed (see Appendix G). These suggested that a single factor was possible, consistent with the high eigenvalues shown in the first factor in all pre-test and post-test results. Two factors were chosen for the final solution, however, because the second factors represented from 5% to 13% of the variance in the

instrument. Retaining factors with eigenvalues in this range is an accepted factor analysis practice (Hatcher & Stepanski, 1994a). The two-factor solution increased the variance accounted for in the B pre-tests results from 81% to 90% and from 70% to 83% in the feedback and training (B) post-test results, which compared favorably with the second factor cumulative eigenvalues of the other two groups (A and BC) in both pre- and post-tests (.93, .86, and .91 and .83, respectively). In addition, this solution was consistent with previous research that shows a high frequency of two factors in multi-rater feedback survey instruments, most often conforming to the relationship and task dichotomy (Craig & Kaiser, 2003). Using a two-factor solution also made these results comparable with past and future research and allowed examination of differences in task and relationship behavior, a well-accepted construct of leadership behavior (Blake & Mouton, 1985; Fiedler, 1971; Hersey & Blanchard, 1988).

Selecting Items for Each Factor. The rotated factor patterns for each exploratory factor analysis (standardized regression coefficients) were examined to identify items that loaded significantly on the two factors. Standard practice in exploratory factor analysis is to delete items that have coefficients less than .40, that do not load clearly on one factor (nearly equal loadings on multiple factors) or shift from one factor in the pre-test results to another factor in the post-test results (Hatcher & Stepanski, 1994a).

The rotated factor pattern for the feedback only (A) group pre-test and post-test exploratory factor analyses are shown in Table 4.6. Based on this factor pattern, the following items were discarded: item 20 (equal loading on both factors in pre-test), item 24 (equal loading on both factors in post-test), and item 28 (loaded on factor 1 in pre-test and

factor 2 in post-test). Items with coefficients of .4 or above are denoted with an asterisk (*). Discarded items are shown in italics.

Table 4.6

Rotated Factor Pattern for Group A

	Pre-test		**Post-test**	
Item	Factor 1	Factor 2	Factor 1	Factor 2
1	56*	30	62*	27
2	26	58*	36	47*
3	118*	-38	113*	-31
4	123*	-43*	131*	-53*
5	87*	-1	93*	-9
6	99*	-12	104*	-17
7	69*	12	65*	23
8	82*	1	85*	2
9	83*	-2	93*	-7
10	78*	3	79*	6
11	83*	-8	75*	3
12	-17	100*	-6	92*
13	-24	110*	-24	110*
14	-29	116*	-18	105*
15	2	81*	-5	90*
16	68*	16	65*	19
17	92*	-8	91*	-5
18	61*	24	68*	16
19	50*	35	72*	10
20	***43****	***44****	51*	36
21	9	76*	30	59*
22	1	75*	8	71*
23	63*	17	76*	6
24	20	70*	***45****	***47****
25	63*	24	63*	24
26	73*	18	76*	15
27	68*	13	81*	-1
28	***31***	***53****	***66****	***15***

The rotated factor pattern for the feedback and training (B) group pre-test and post-test exploratory factor analyses are shown in Table 4.7. Based on this factor pattern, the following items were discarded: item 18 (loaded on factor 1 in pre-test and factor 2 in post-test), item 23 (loaded on factor 2 in pre-test and factor 1 in post-test), item 25 (loaded on factor 2 in pre-test and factor 1 in post-test), item 27 (equal loading on both factors in post-test), and item 28 (loaded on factor 1 in pre-test and factor 2 in pre-test). Items with coefficients of .4 or above are shown with an asterisk (*). Discarded items are shown in italics.

Table 4.7.

Rotated Factor Pattern for Group B

	Pre-test		**Post-test**	
Item	Factor 1	Factor 2	Factor 1	Factor 2
1	69*	6	63*	5
2	34	43*	32	41*
3	92*	-4	109*	-39
4	115*	-37	94*	-35
5	102*	-22	89*	-17
6	95*	-10	104*	-25
7	70*	19	71*	10
8	65*	23	74*	5
9	67*	20	67*	9
10	63*	24	79*	2
11	47*	34	69*	5
12	-1	85*	2	75*
13	-27	116*	-20	102*
14	-15	102*	-16	98*
15	31	59*	22	54*
16	72*	11	59*	17
17	100*	-21	72*	-2
18	67*	14	***29***	***42****
19	69*	9	53*	28
20	71*	14	62*	15
21	-4	91*	10	64*
22	-25	106*	-19	85*
23	3	76*	***54****	***24***
24	31	59*	36	45*
25	20	70*	***55****	***33***
26	68*	24	77*	12
27	***43****	***41****	58*	8
28	-11	80*	14	55*

The rotated factor pattern for the feedback and training and the feedback, training and coaching (BC) groups pre-test and post-test exploratory factor analyses are shown in Table 4.8. Based on this factor pattern, the following items were discarded: item 18 (loaded on factor 1 in pre-test and factor 2 in post-test), item 23 (loaded on factor 2 in pre-test and factor 1 in post-test), item 24 (equal loading on both factors in post-test), item 25 (loaded on factor 2 in pre-test and factor 1 in post-test), and item 27 (loaded on factor 2 in pre-test and factor 1 in post-test). Items with coefficients .4 or above are shown with an asterisk (*). Discarded items are shown in italics.

Table 4.8

Rotated Factor Pattern for Group BC

	Pre-test		**Post-test**	
Item	Factor 1	Factor 2	Factor 2	Factor 2
1	73*	1	52*	14
2	34	42*	36	40*
3	72*	2	111*	-40
4	114*	-38	100*	-33
5	103*	-25	93*	-15
6	96*	-12	104*	-21
7	69*	16	79*	2
8	61*	24	84*	-2
9	79*	4	75*	5
10	68*	15	81*	1
11	57*	22	73*	3
12	-10	93*	-8	82*
13	-27	114*	-19	102*
14	-20	105*	-15	99*
15	10	76*	24	59*
16	45*	35	58*	20
17	96*	-18	77*	0
18	64*	17	***31***	***46****
19	76*	3	65*	19
20	76*	9	68*	14
21	0	85*	14	63*
22	-21	100*	-15	77*
23	16	63*	***71****	***9***
24	30	58*	**45***	**42***
25	22	68*	**67***	**21**
26	70*	11	80*	13
27	39	43*	***53****	***18***
28	3	72*	30	50*

In summary, seven items were deleted. Items were deleted for the following reasons: did not load significantly on any factors (18), loaded evenly with a low score (41-45) on one or more factors (20, 24, 27), or items loaded on a different factor in the post-test (18, 23, 25, 27, 28). These are shown in Table 4.9.

Table 4.9

Summary of Deleted Items

Deleted Items	Rationale
18. Deals well with uncertainty and ambiguity	Shifted from F1 to F2 in B post-test; did not load significantly on either factor in BC post-test
20. Considers alternatives/consequences before making decisions	Loaded evenly and low on A pre-test
23. Understands and practices appropriate ethics	Shifted from F2 to F1 in B and BC post-tests
24. Provides purpose, values, vision	Loaded evenly and low on A all post-tests (high 45, difference of 2 points in two tests)
25. Takes responsibility for his/her actions	Shifted from F2 to F1 in B and BC post-tests
27. Effectively deals with stress and frustration	Loaded evenly and low (43-41) on B pre-test; shifted from F2 to F1 in BC post-test
28. Projects poise and confidence	Shifted from F2 to F1 in A post-test

The resulting scales are shown in Table 4.10. An examination of the items that loaded on the two factors showed that the items on the first factor appeared to reflect relationship behavior (Blake & Mouton, 1985; Fiedler, 1971; Hersey & Blanchard, 1988). The items that loaded on the second factor appeared more consistent with task-oriented behavior (Blake & Mouton, 1985; Fiedler, 1971; Hersey & Blanchard, 1988). There are 14 items on the first or relationship factor and seven on the second or task factor. These are shown in Table 4.10. Mean scores between pre- and post-test measures for the relationship scale and the task scale were analyzed separately.

Table 4.10

Relationship and Task Factors

Factor 1 Relationship	Factor 2 Task
1. Demonstrates effective listening and questioning skills	2. Effectively articulates information
3. Sensitive to needs of others	12. Seeks new challenges and responsibilities
4. Shows tact in dealing with others	13. Self-starter and finisher
5. Successfully wins cooperation of others	14. Persistent in overcoming obstacles
6. Earns trust and loyalty	15. Develops innovative ideas and approaches that work
7. Willingly helps and provides resources to others	21. Willing/able to make tough decisions when required
8. Acts as a member of the team	22. Does not procrastinate
9. Places the team before personal interest	
10. Gives credit to others	
11. Avoids "playing politics"	
16. Open to new ideas and approaches	
17. Flexible in dealing with different viewpoints and styles	
19. Consults with others in dealing with difficult situations	
26. Earns trust and loyalty, inspires others to greater performance	

In addition, three-factor solutions were also examined as a potential factor structure, but they were rejected because the incremental increase in cumulative eigenvalues ranged from .02 to .08, which meant that the three-factor solution would have resulted in deleting 13 items from the 28-item instrument.

Constructing Factor Scores. Factor scores were constructed by calculating the mean score for each item on the factor across all respondents for each participant in the pre-test, summing them, and calculating the mean factor score. In multiple analysis of covariance, differences between groups are tested by examining the interaction effect of the group and

the factor scores at time 1 and time 2. This is adjusted by the appropriate covariate values (Rawlings et al., 1998). The factor-based scale method was used instead of estimated factor scoring methodology because the two methods correlate to .99 (Hatcher & Stepanski, 1994a).

This process resulted in two dependent variables for further testing: change in relationship scores over time and change in task scores over time.

Testing the Research Question

The following section provides the results of the statistical analyses used to test this study's research question: Are there significant differences in the change scores in leadership behavior for each of the three following groups:

1. leaders who received 360-degree feedback only on their leadership behavior from their superiors, peers, and subordinates;
2. leaders who received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates, and who also attended a five-day leadership training program designed to promote leadership behavior change;
3. leaders who: (a) received 360-degree feedback on their leadership behavior from superiors, peers, and subordinates; (b) attended a five-day leadership training program designed to promote leadership behavior change; and (c) received executive coaching, which consisted of 10 to 20 hours of one-on-one coach interaction with the client, or who received peer coaching, which consisted of 16 hours of group interaction with peers and a trained consultant?

MANCOVA for the Relationship Factor Scores. A multiple analysis of covariance was calculated to test whether there were significant differences in the change in relationship scores between the three groups. Age of participants and time between administration of pre-test and post-test were used as covariates. The mean scores and standard deviations (Std Dev) of each of the three groups—feedback only (A), feedback and training (B) and feedback, training, and coaching (C)—for the difference in relationship scores and the covariates are shown in Table 4.11.

Table 4.11

Means and Standard Deviations of Relationship Factor Scores and Covariates

		Relationship Difference Score		**Time (Covariate)**		**Age (Covariate)**	
Group	N	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
A	71	0.1385237	.6183652	208.5633803	25.6886481	40.4285714	7.6282682
B	21	0.3103907	.9867842	580.1904762	270.8761007	39.5714286	7.1033191
C	9	-.1661773	.5682983	555.5555556	72.6853091	42.1111111	5.8404718

The mean for relationship difference scores in the feedback only (A) group was .1385. The mean for relationship difference score for the feedback and training (B) group was more than twice that: .3104. The mean for relationship difference score for the feedback, training, and coaching group (C) actually declined by -.1662. The mean score for the covariate of time was 209 for the A group, 580 for the B group, and 556 for the C group. The mean score for the covariate age was 40 for the A and B groups and 42 for the C group.

The results of the multiple analysis of covariance are shown in Table 4.12. The results of the analysis indicate that there were no significant differences between the change in relationship scores over time as measured by the interaction effect of group and test, when

adjusted for the covariates of time and participant age, $F = 0.35$, $p > .7056$. For the effect of group alone [feedback only (A), feedback and training (B), or feedback, training, and coaching (C)], the analysis failed to show significant differences, $F = 1.12$, $p > .32$; the effect of time was also not significant, $F = .35$, $p > .56$. The effect for age was significant, $F = 6.85$, $p < .01$. The Pearson product moment correlation between participants' ages and the relationship difference scores was .14, which was significant at $p < .04$.

Table 4.12

Multiple Analysis of Covariance of Relationship Factor Scores

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	*F* Value	Pr > *F*
Group	2	190	1.12	0.3289
Test	1	190	0.13	0.7220
Group*test	2	190	0.35	0.5569
Time	1	190	0.35	0.5569
Partage	1	190	6.85	0.0096

MANCOVA for the Task Factor Scores. A multiple analysis of covariance was calculated to test whether there were significant differences in the task scores over time among the three groups. Age of participants and time between the administration of the pre-test and post-test were used as covariates. The mean difference scores and standard deviations (Std Dev) of each of the three groups [feedback only (A), feedback and training (B), and feedback, training, and coaching (C)] are shown in Table 4.13.

Table 4.13

Means and Standard Deviations of Task Factor Scores and Covariates

		Task Difference Score		Time (Covariate)		Age (Covariate)	
Group	N	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
A	71	0.1236335	.6314441	208.5633803	25.6886481	40.4285714	7.6282682
B	21	0.3652955	.8919184	580.1904762	270.8761007	39.5714286	7.1033191
C	9	-.2146709	.3781750	555.5555556	272.6853091	42.1111111	5.8404718

The mean for the task difference score for the feedback only (A) group was .1236. The mean for the task difference score for the feedback and training (B) group was three times that: .3653. The mean for the task difference score for the feedback, training and coaching group actually declined by -.2147. The mean score for the covariate of time was 209 for the A group, 580 for the B group, and 556 for the C group. The mean score for the covariate age was 40 for the A and B groups and 42 for the C group.

The results of the multiple analysis of covariance are shown in Table 4.14.

Table 4.14

Multiple Analysis of Covariance of Task Factor Scores

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	*F* Value	Pr > *F*
Group	2	190	2.51	0.0843
Test	1	190	0.05	0.8174
Group*test	2	190	0.46	0.6312
Time	1	190	2.19	0.1405
Partage	1	190	3.18	0.0762

The results of the analysis indicate that there were no significant differences among

the three groups for change in task factor scores over time as measured by the interaction effect of group and test, when adjusted for the covariates of time and participant age, $F = 0.46$, $p > .6312$. For the effect of group alone [feedback only (A), feedback and training (B), or feedback, training, and coaching (C)], the analysis failed to show significant differences also, $F = 2.51$, $p > .08$; the effect of time was also not significant, $F = 2.19$, $p > .14$. The effect for age was also not significant, $F = 3.18$, $p > .08$ in the covariate model; however, when time was deleted from the model, age was a significant covariate with $F = 4.35$, $p < .04$. The Pearson product moment correlation of age and task factor scores was .32, which is significant, $p < .0001$.

Because the effect of group approached the level of significance, a two-way analysis of variance with interaction effect (Agresti & Finlay, 1997) but no covariates was conducted to assess differences among the three groups solely based on group. The results of this analysis are shown in Table 4.15.

Table 4.15

Two-Way ANOVA for Group on Task Factor Scores

Type 3 Tests of Fixed Effects

Source	Num DF	Den DF	*F* Value	Pr > *F*
Group	2	195	12.05	0.0001
Test	1	195	0.07	0.7958
Group*test	2	195	0.46	0.8133

The main effect of group was significantly different across the three groups, $F = 12.05$, $p < .0001$. Effect tests were conducted to determine if there were pre-existing differences among the three groups on both pre-test scores and post-test scores. Effect tests

provide an F ratio fro each effect and test the null hypothesis that the true value of the parameter is zero (T. Johnson & Berk, 2000). The results of these tests are shown in Table 4.16.

Table 4.16

Test of Effect Slices for Pre-test and Post-test Scores by Group

Type 3 Tests of Effect Slices

Source	Num DF	Den DF	*F* Value	Pr > *F*
Pre-test	2	195	7.24	0.0009
Post-test	2	195	5.27	0.0059

The test results indicated that there were significant differences among the three pre-test groups ($F = 7.24$, $p < .001$) and among the post-test groups ($F = 5.27$, $p < .001$).

Differences Among Pre-test Scores on the Task Factor. The least squares means and standard errors of pre-test scores on the task factor are shown for the three groups in Table 4.17. Least squares means were adjusted for the covariates used in the study.

Table 4.17

Least Squares Means and Standard Errors of Task Pre-test Scores

Task Pre-test Scores

Group	N	Mean	Std Err
A	71	7.5821	0.09346
B	21	8.0788	0.1773
C	9	8.4902	0.2644

The Tukey-Kramer (Agresti & Finlay, 1997) Test of Differences of Least Squares Means was performed to determine which of the pre-test groups differed significantly from

each other. Results of this test are shown in Table 4.18. The results showed that the feedback only (A) and feedback, training, and coaching (C) groups differed significantly from each other, but the feedback and training (B) group did not differ significantly from the other groups.

Table 4.18

Tukey-Kramer Test of Least Squares Means for Pre-test Scores on the Task Factor

Group Comparison	Estimate	Standard Error	Adjusted *p* value
A - B	-0.4967	0.2005	0.1358
B - C	-0.4113	0.3183	0.7891
A - C	-0.9081	0.2804	0.0175***

Note: *** denotes comparisons that are significant at $p > .05$

Differences Among Post-test Scores on the Task Factor

The least squares means and standard errors of pre-test scores on the task factor are shown for the three groups in Table 4.19. Least squares means were adjusted for the covariates used in the study.

Table 4.19

Least Squares Means and Standard Errors of Task Post-test Scores

	Task Post-test Scores		
Group	N	Mean	Std Err
A	71	7.7138	0.09412
B	21	8.2835	0.1773
C	9	8.2755	0.2644

The Tukey-Kramer (Agresti & Finlay, 1997) Test of Differences of Least Squares Means was performed to determine which of the post-test groups differed from each other. Results of this test are shown in Table 4.20.

Table 4.20

Tukey-Kramer Test of Least Squares Means for Post-test Scores on the Task Factor

Group Comparison	**Estimate**	**Standard Error**	**Adjusted *p* value**
A - B	-0.5697	0.2806	0.0558
B - C	0.008015	0.3183	1.000
A - C	-0.5617	0.2806	0.3451

Note: *** denotes comparisons that are significant at $p > .05$

The results failed to show any significant difference among the groups, although a trend may be present between the feedback only (A) and the feedback and training (B) groups.

Two-Way ANOVA for Gender

Demographic data on the groups showed that they differed considerably with regard to gender. The feedback only (A) group was 36% male. The feedback and training (B) group was 64% male. The feedback, training, and coaching (C) group was 67% male. In order to determine whether gender was significantly related to relationship and task difference scores, factorial analyses of variance with two between-group factors were performed (Hatcher & Stepanski, 1994b). The results of the analysis for the relationship difference scores are shown in Table 4.21.

Table 4.21

Two-Way ANOVA for Gender on Relationship Difference Scores

Source	DF	Type III SS	Mean Square	*F* Value	Pr > *F*
Group	2	1.93875816	0.96937908	1.92	0.1516
Gender	1	0.00437681	0.00437681	0.01	0.9259
Gender*Group	2	0.83841979	0.41920990	0.83	0.4382

The analysis for relationship difference scores failed to show any significant effects among groups, between genders, or among the interactions of the two. The *F* value for the main effect of group was 1.92, with $p > .15$. The *F* value for the main effect of gender was 0.01, with $p > .93$. The *F* value for the interaction effect of group and gender was 0.83, with $p > .44$.

The results of the analysis for the task difference scores are shown in Table 4.22.

Table 4.22

Two-Way ANOVA for Gender on Task Difference Scores

Source	DF	Type III SS	Mean Square	*F* Value	Pr > *F*
Group	2	1.76890968	0.88445484	1.88	0.1587
Gender	1	0.07220495	0.07220495	0.15	0.6963
Gender*Group	2	0.19518937	0.09759468	0.21	0.8133

The analysis for task difference scores failed to show any significant effects among groups, between genders, or among the interactions of the two. The *F* value for the main effect of group was 1.88, with $p > .16$. The *F* value for the main effect of gender was 0.15, with $p > .70$. The *F* value for the interaction effect of group and gender was 0.21, with $p > .81$.

Paired Sample t *Tests Within each Group*

To determine whether each of the groups experienced significant change from pre-test to post-test, results were analyzed by group using paired-sample *t* tests (Hatcher & Stepanski, 1994b). The results of these tests are shown in Table 4.23.

Table 4.23

Paired Sample t *Tests by Group*

		Relationship Difference Scores			**Task Difference Scores**		
Group	N	Mean	Std Err	$p > \underline{t}$	Mean	Std Err	$p > t$
A	71	0.1385237	.0733865	0.0632	0.1236335	.0749386	0.1035
B	21	0.3103907	.2153340	0.1649	0.3652955	.1946323	0.0752
C	9	-0.1661773	.1894328	0.4059	-0.1260583	.1260583	0.1270

For the relationship factor in the feedback only group (A), the test failed to show a significant difference between pre-test and post-test scores, $p > .06$. For the task factor in the feedback only group (A), the test failed to show a significant difference between pre-test and post-test scores, $p > .10$.

For the relationship factor in the feedback and training group (B), the test failed to show a significant different between pre-test and post-test scores, $p > .17$. For the task factor in the feedback and training group (B), the test failed to show a significant difference between pre-test and post-test scores, $p > .08$.

For the relationship factor in the feedback, training ,and coaching group (C), the test failed to show a significant difference between pre-test and post-test scores, $p > .41$. For the task factor in the feedback, training, and coaching group (C), the test failed to show a significant difference between pre-test and post-test scores, $p > .13$.

Summary of Results

Pre-test and post-test responses to the 28-item proprietary leadership effectiveness survey were subjected to an exploratory factor analysis using squared multiple correlations as prior communality estimates. Pre-test and post-test results were analyzed to ensure that they

were equivalent. The maximum likelihood factor method was used to extract the factors, and this was followed by a promax (oblique) rotation. Scree tests, preliminary eigenvalues and an analysis of conceptual meaning of the items suggested two meaningful factors that were retained for rotation.

In interpreting the rotated factor patterns, an item was said to load on a given factor if the factor loading was .40 or greater for that factor and was less than .40 for the other factor. Items were also discarded if they shifted factor loading between pre-test and post-test exploratory factor analyses. Using these criteria, 14 items were found to load on the first factor, which was labeled the relationship factor. Seven items loaded on the second factor, which was labeled the task factor. Seven items were discarded from further analyses. Factor-based difference scores between pre-test and post-test were constructed for the relationship and task factors. These were analyzed separately in all statistical tests of difference.

There were no significant differences between the change in relationship factor scores over time as measured by the interaction effect of group and test, when adjusted for the covariates of time and participant age. Two covariates, age and time between pre-test and post-test administration, also failed to show significant differences across all three groups.

There were no significant differences among the three groups for change in task factor scores over time as measured by the interaction effect of group and test, when adjusted for the covariates of time and participant age; however, the effect for group approached significance. Effect tests found highly significant differences in both pre-test and post-test scores for the task factor.

Tests of pre-test scores failed to show any significant pre-existing differences among the three groups for the relationship factor. Tests did show a significant difference between

the feedback only (A) group and the feedback, training, and coaching (C) group on the pre-test task factor.

Tests of post-test scores failed to show any significant differences among the three groups for the relationship factor. Tests did show a trend toward a significant difference between the feedback only (A) group and the feedback and training (B) group on the post-test task factor.

There was also a weak, but significant, positive correlation between both relationship and task factor scores and participant age.

Tests for differences among the groups due to gender also failed to show significant differences. Finally, tests for differences between pre-tests and post-tests within each of the three groups failed to show significant differences, although possible trends are discussed in Chapter 5.

Chapter 5: Discussion

The primary purpose of this study was to determine whether there were significant differences in changes in leadership behavior due to feedback, training, and coaching. There were three groups in the study: participants who received feedback only (A), participants who received feedback and training (B), and participants who received feedback, training, and coaching (C). Pre-test and post-test responses were subjected to exploratory factor analyses, and two factors emerged: a relationship factor and a task factor. The change over time in their scores between pre-test and post-test administration of 360-degree feedback surveys on both relationship and task factors were used as the dependent variables.

This chapter begins with a discussion of the findings of analyses directly related to the research question and other analyses done to better understand these findings. Implications of findings for theory and practice are also discussed in this chapter, followed by the study's limitations. The limitations address the issue of attrition (participants who failed to complete the post-test) in all groups, the small sample size of the feedback, training, and coaching (C) group, the possible impact of adding peer-coaching clients to that group, and a comparison of participants in that group with those who chose not to participate in the study. Finally, recommendations for future research are made.

Differences in Leadership Behavior Change Across Groups

Differences in relationship factor scores across groups. For the change over time in scores on the relationship factor, there were no significant differences among the three groups when adjusted for the covariates of age and time. Reasons that no significant differences were found may include the small sample sizes in two of the groups, the sensitivity of the relationship scale items, or many of the other factors that make it difficult to

measure behavioral change resulting from feedback, training, or coaching. Feedback does not automatically result in behavior change (Brutus, Fleenor, & London, 1998; Kaplan, Drath, & Kofodimos, 1987; London & Smither, 1995). Kluger and DeNisi (1966) reviewed more than 600 feedback studies and found that one third of feedback interventions resulted in negative changes in performance, one third showed no change in performance, and one-third did have a positive impact on performance.

Dalton and Hollenbeck (2001) cite a meta-analysis done by Hellervik, Hazucha, and Schneider (1992) showing that the effectiveness of psychotherapy is .78 of a standard deviation, while the effectiveness of management development training is only .43. They also point out the following estimates of the amounts of behavior change after coaching, which might be expected to yield greater change than feedback (Dalton & Hollenbeck, 2001): Baird and Bolton (1999) estimate that 90% of participants will have increased self-awareness; 70% will make some attempt to try out new behaviors; and only 20% will make the new behavior a permanent part of their behavioral repertoire. These estimates are consistent with self-report data collected by McCauley and Hughes-James (1994) following a program that contained feedback, training, and coaching. They found that 58% reported considerable change, 29% reported little change, and 13% reported no change. Conger (1992) reported on follow-up after leadership development programs that included feedback. He reported that 10–20% showed no change and little enhanced self-awareness; 30–40% showed expanded self awareness; 25–30% showed some positive behavior change; and only 10% showed significant behavior change (Conger, 1992). The levels of intensity of these interventions, progressing from feedback alone, to feedback and training, up to feedback, training, and

coaching, may be a partial explanation for the variance in these figures (Dalton & Hollenbeck, 2001).

Differences in pre-test task factor scores across groups. For the change over time in scores on the task factor, there were no significant differences among the three groups when adjusted for the covariates of age and time; however, the main effect for group showed a highly significant difference among the three groups, which was confirmed by a test of effect slices. A test of differences of least squares means found a significant difference between the feedback only (A) group and the feedback, training, and coaching (C) group on the pre-test task factor. There is almost a full point of difference between the two groups (7.58–8.49). Participants in the feedback, training, and coaching (C) group had significantly higher task factor pre-test scores than participants in the feedback only (A) group. Participants in the feedback and training group (B) had pre-test scores of 7.97 on the task factor, approximately halfway between the other two groups. Although this difference is not significant, it appears that there may be a trend for the feedback and training (B) group to share similar characteristics with the feedback, training, and coaching (C) group.

One possible explanation for this difference among the groups is that participants in the C group may have been selected or self-selected themselves into the program because of their more highly developed task-oriented leadership behaviors. Many scholars agree that coaching participants are more likely to be at risk for derailment (Day, 2000; Dyer, 2002; Hagberg, 1996; Hellervik et al., 1992). The most commonly reported causes of derailment include arrogance, abrasiveness, and poor interpersonal relations (Lombardo et al., 1994; McCall & Lombardo, 1983). A common pattern observed in executives who derail is that strengths become weaknesses if used in excess or not modified in relation to changing

demands (Lombardo & Eichinger, 1989; McCall & Lombardo, 1983; McCall, Lombardo, & Morrison, 1988). Strengths of those who derail include independence, being less affiliative, being assertive, being introspective, having high energy and initiative, having the ability to solve problems well and quickly, and having the ability to take charge (Bass, 1985; Bray, Campbell, & Grant, 1974; Lombardo & Eichinger, 1989). A number of scholars describe the derailment process this way: "People who are successful early in their careers appear to be proficient in task-based leadership but are presented with a challenge when job demands begin to require a balance with a more relationship-oriented style" (Leslie & Van Velsor, 1996, p. 16). Many of these characteristics align with the behaviors that loaded on the task factor in this study. These items are shown in Table 5.1.

Table 5.1

Task Factor Items

Effectively articulates information
Seeks new challenges and responsibilities
Self-starter and finisher
Persistent in overcoming obstacles
Develops innovative ideas and approaches that work
Willing/able to make tough decisions when required
Does not procrastinate

It is also possible that this strength in task-oriented leadership behavior declined for these participants at least in part because they may have been focusing on new relationship-oriented behaviors or were even consciously trying to disengage from over-reliance on task-oriented behaviors. Goal-setting and learning strategies participants developed with their coaches may have included such strategies (Lombardo & Eichinger, 1989; Ting, S., & Hart,

2004). A downward shift on some scales was identified in an earlier study (McCauley & Hughes-James, 1994) and explained in a similar fashion.

Differences in post-test task factor scores across groups. Tests of differences of least squares means showed a difference closely approaching the level of significance ($p = .0558$) between the feedback only (A) group and the feedback and training (B) group on the post-test task factor. There is a large difference in absolute scores between the A group (7.7) and the B group (8.28). It is also noted that there is no significant difference between the B group (8.2835) and the C group (8.2755). It is helpful to examine the mean difference scores for both pre-tests and post-tests on the task factor across all three groups in order to analyze trends. These are shown in Table 5.2.

Table 5.2

Least Squares Means and Standard Errors of Task Pre-test and Post-test Scores

		Task Pre-test		**Task Post-test**	
Group	N	Mean	Std Err	Mean	Std Err
A	71	7.5821	0.09346	7.7138	0.09412
B	21	8.0788	0.1773	8.2835	0.1773
C	9	8.4902	0.2644	8.2755	0.2644

There is a significant difference between pre-test means between groups A and C. This is a difference of almost one full point (7.58–8.49). In post-test means, there is a large but non-significant difference between groups A and B. This is a difference of approximately three fourths of a point. There is virtually no difference between the means of group B and C. Group B changed the most in absolute numbers. This might suggest that there is an incremental effect of training and feedback over feedback alone, at least on the task factor. Because of the actual decline in post-test scores for the feedback, training, and coaching (C)

group, similar patterns of differences were not found. Potentially different dynamics of the change process within the C group may be due to derailment propensities, as discussed above.

Decline in Post-test Scores for Feedback, Training, and Coaching (C) Group

The decline (negative gain scores) in leadership behavior change scores on both relationship (-.17) and task (-.21) factors for the feedback, training, and coaching (C) group was unexpected. Other studies have reported some declines in behavior change scores after similar interventions. As cited earlier, one third of more than 600 feedback studies resulted in negative changes in performance (Kluger & DeNisi, 1996). McCauley and Hughes-James (1994) reported 36% downward shifts in ratings across all 16 scales used in their study. They attributed some of the downward shifts to participants becoming more reflective and planful as a result of working through a conscious change process; some of the downward shift was also attributed to external events that were stressful for their participants; but the researchers did acknowledge that the remaining downward shifts were puzzling. An examination of individual means in the feedback, training, and coaching group (C) in this study showed that 56% had downward shifts in their ratings. Four of the nine participants had downward shifts on the relationship factor and six of the nine on the task factor.

When the feedback received from the feedback report in a training program or from a coach is discrepant with the individual's self-concept, the individual may react by rejecting the data through the process of denial, which decreases the likelihood of behavior change (Dalton, 1997). In this study, self-evaluations from participants were not collected, so it was impossible to assess who might have received discrepant information. However, researchers on feedback estimate that approximately one-third of participants are likely to rate

themselves higher than their direct reports will rate them (Fleenor & Prince, 1997). Therefore, it seems reasonable to expect that at least one third of participants in this study received discrepant information and may have rejected it, resulting in a deterioration of their skills along the task dimension.

In a landmark study at GE (Meyer, Kaye, & French, 1965), researchers found that managers who received negative feedback from their bosses that was tied directly to salary decisions actually showed a decrease in performance. It was impossible to assess the extent to which participants may have perceived that the feedback they received could affect salary decisions or career progress. However, a number of researchers (Day, 2000; Dyer, 2002; Hagberg, 1996; Hellervik et al., 1992) contend that executive coaching tends to be reserved for senior executives most at risk for derailment. If this is true, it is likely that more participants in this group may have perceived relatively immediate negative career consequences associated with their efforts to change. It is likely that some, if not all, of these participants discussed with their bosses or others in the organization, such as human resources managers, why they were chosen to attend this development program. This is a recommended common practice prior to participating in a coaching program (Ting et al., 2004; Witherspoon & White, 1997).

The fact that pre-test scores of the feedback, training, and coaching (C) group were actually significantly higher than the feedback only (A) group might also suggest that the actual decline in those post-test scores could be attributed to regression to the mean. Walker and Smither (1999) found that for managers who received the most favorable ratings in pre-test feedback, all of their deterioration in scores in the post-test could be attributed to regression to the mean.

Differences Between Pre-test and Post-test Scores Across Groups

One final question remains about the behavior changes in the three groups: Did any of the groups experience significant change from pre-test to post-test? Paired sample *t* tests failed to show significant differences between pre-test and post-test means on both relationship and task factors; however, an examination of *t* values suggests that a trend might be evident. Table 5.3 shows the means, standard errors, and probabilities for task and relationship means across the groups.

Table 5.3

Paired Sample t *Tests by Group*

		Relationship Difference Scores			**Task Difference Scores**		
Group	N	Mean	Std Err	$p > \underline{t}$	Mean	Std Err	$p > \underline{t}$
A	71	0.1385237	.0733865	0.0632	0.1236335	.0749386	0.1035
B	21	0.3103907	.2153340	0.1649	0.3652955	.1946323	0.0752
C	9	-0.1661773	.1894328	0.4059	-0.1260583	.1260583	0.1270

For the feedback only (A) group, the difference between pre-test and post-test scores on the relationship factor approached the level of significance, at $p > .06$. On the task factor, the difference was not as great, at $p > .10$. The failure to detect any differences due to feedback alone might seem surprising in light of the previously cited evidence that feedback alone can account for significant behavioral changes (Atwater et al., 1995; Bernardin et al., 1995; Hazucha et al., 1992; Hegarty, 1974; J. W. Johnson & Ferstl, 1999; Reilly et al., 1996; Smither & Walker, 1995; Walker & Smither, 1999). Most of the studies cited earlier were conducted in a performance appraisal context in which participants were required to share their feedback results with their bosses. This accountability for performance improvement has been identified as an intervening variable that may provide significantly more motivation

to change than feedback received in a purely developmental framework in which feedback ratings are confidential (Dalton & Hollenbeck, 2001; London, 2001; London, Smither, & Adsit, 1997). This study was conducted in a developmental context in which participants were free to share their feedback with their bosses but not required or encouraged to do so. In a post-test questionnaire, participants in the feedback only (A) group reported that 53% did share their results with their bosses, compared to 67% and 86%, respectively, in the feedback and training (B) and feedback, training, and coaching (C) groups. Lower levels of accountability may have contributed to changes being nonsignificant in the feedback only (A) group.

The pattern evident in the feedback only (A) group (stronger relationship factor differences vs. task factor differences) is the opposite of that shown in the other two groups. In both the feedback and training (B) and the feedback, training, and coaching (C) groups, task factor differences scores had lower probabilities (.08 and .13, respectively) than relationship factor differences (.17 and .41, respectively). The number of participants in these groups was very small (21 and 9, respectively) and may have affected results; however, it does appear that training or coaching may affect outcomes differently at least on the task factor.

Summary of Findings

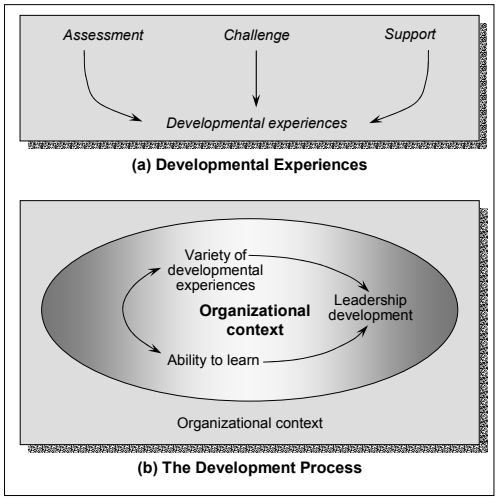
It does appear from the analyses of both pre-test and post-test scores that the focus of behavior that participants try to change may be affected by training or coaching. In both cases, there is evidence that task-oriented behaviors change more after training than relationship behaviors. It may be that task-oriented behaviors change first, are easier to change, or are easier for feedback respondents to recognize. In this study, at least, feedback,

training, and coaching (C) caused performance of task-oriented behaviors to decline, but not to a significant degree. The participants in this group seemed to differ from participants in the other groups based on their significantly higher pre-test scores on the task factor. The dynamics of change for this population, whose members are more likely to be at risk for derailment, may well be different from the dynamics for members of other groups, as many previously cited researchers have suggested (Leslie & Van Velsor, 1996; Lombardo & Eichinger, 1989; McCall & Lombardo, 1983). It also appears that training or coaching may affect change outcomes differently from the way that feedback alone does.

Implications for Theory and Practice

Implications for Theory. The most comprehensive model or theory of leader development was published by the Center for Creative Leadership in 1998 and revised in 2004. The model, shown in Figure 5.1, has two parts. In part (a), Developmental Experiences, the model points out that developmental experiences of any kind (feedback, training, coaching, projects, keeping a journal, etc.) are effective as experiences that help people learn and change because of three elements: assessment, challenge, and support. Experiences that add additional levels of each of these elements can be even more developmental for individuals. This aspect of the model may be exceptionally helpful to those who design feedback, training, and coaching programs.

Figure 5.1. Center for Creative Leadership leader development model.



From "Introduction: Our view of leadership development," by E. Van Velsor and C. D. McCauley in *The Center for Creative Leadership handbook of leadership development* , by C. D. McCauley and E. Van Velsor (eds.), p. 4. Copyright 2004 Jossey-Bass.

pds

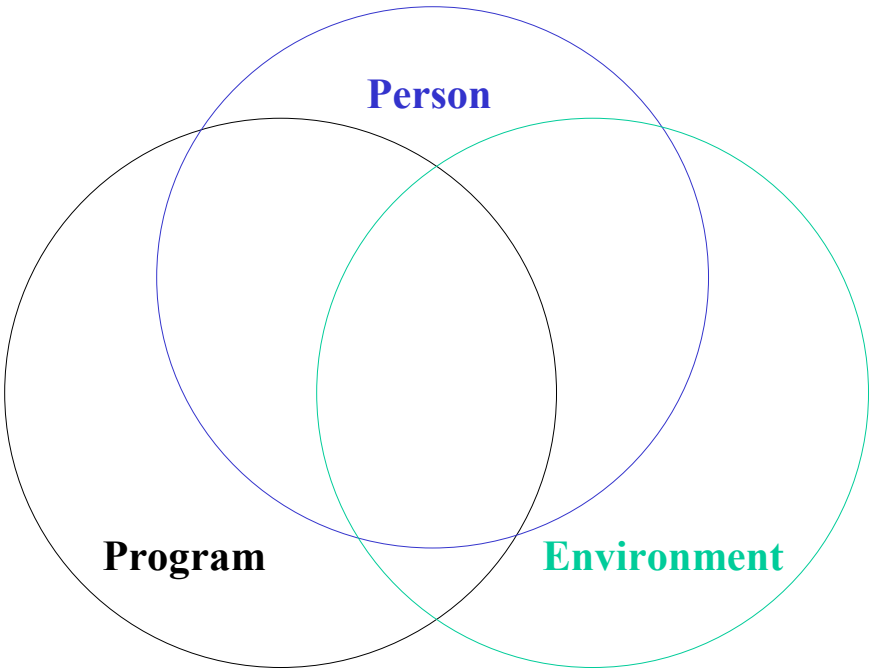
Part (b) of the model, The Developmental Process, may be more helpful in explaining the results of this study. In part (b), there are three main principles: A variety of developmental experiences are necessary for development; an individual's ability to learn is important and can be enhanced by developmental experiences; and a leader development program is embedded in an organizational context that must support the individual and the development process (McCauley & Van Velsor, 2004).

This study focused on variables that described differences in developmental experiences. No variables related to participants' different abilities to learn or organizational context were explored, although it does appear that preexisting differences in potential derailment factors may well have affected results. It is also interesting to note that participant age was significantly positively correlated with both task and relationship factor scores. For purposes of this study, it may be more helpful to reframe the Center for Creative Leadership

model in a way similar to Baldwin and Ford's multistage model of transfer of training. In that model, they identify three components that drive transfer of training: trainee characteristics, training design, and the work environment (Baldwin & Ford, 1988).

Leadership behavior change can be viewed as a combination of three different components: characteristics of the person, the program, and the environment. This can be conceptualized as shown in Figure 5.2.

Figure 5.2. Three-stage model of leadership behavior change.



McCauley and Hughes-James (1994) noted that there are variations in program outcomes based on the psychological readiness of the person participating in it. These factors include motivation, capability, willingness, and commitment to change (Ting et al., 2004). More specifically, these factors can be defined as intrapersonal intelligence, cognitive complexity, metacognition, self-esteem, self-efficacy, openness to experience, and conscientiousness (Van Velsor, Moxley, & Bunker, 2004). Many of these characteristics are

well established empirically; however, it may be helpful to define intrapersonal intelligence, which is less well-known. Intrapersonal intelligence is defined as "… the ability to form an accurate model of oneself and to be able to use that model to operate effectively in everyday life" (Van Velsor et al., 2004, p. 214). Little research has been reported on how individual differences among participants interact with differences in program intervention, such as the three conditions used in this study, to effect changes in leadership behavior. It seems likely that understanding how these factors affect outcomes would enable better prediction of behavioral change. Such an examination would allow us to better understand how potential derailment characteristics, for example, affect behavioral outcomes.

Program characteristics include the degree of assessment, challenge, and support present in the intervention (McCauley & Van Velsor, 2004), as well as differences in delivery methods, such as feedback only; feedback and training; feedback, training, and coaching; or other combinations, such as feedback and coaching, which have been implemented more recently (Goldsmith, 2000; Smither et al., 2003). It seems likely that these characteristics interact with personal characteristics in complex ways to produce behavior change.

The environment of a participant in a feedback-based program aimed at improving leadership effectiveness includes primarily the people in it. People in the organization can provide support for the changes the participant desires to make (Van Velsor et al., 2004). One relationship that has critical importance is the relationship with the participant's boss or manager (Dalton & Hollenbeck, 1996). The role of the manager includes: linking behavior change of the participant to business results; providing developmental experiences, such as projects and assignments; providing support and feedback; and providing organizational

support (Dalton & Hollenbeck, 1996). There is emerging empirical evidence suggesting that how participants interact with their bosses can impact the degree of behavior change they are able to make and sustain. Feedback alone when delivered in a performance-appraisal context in which participants review their feedback with their bosses and are held accountable for planning and implementing improvements has been shown to account for significant behavioral changes (Atwater et al., 1995; Bernardin et al., 1995; Hazucha et al., 1992; Hegarty, 1974; J. W. Johnson & Ferstl, 1999; Reilly et al., 1996; Smither & Walker, 1995; Walker & Smither, 1999).

The technique of involving others, particularly direct reports, has received recent focus from other researchers. Walker and Smither (1999) found that managers who met with their subordinates to discuss their feedback and plans for improvement improved more than those who did not. They also found that managers improved more in years when they discussed the prior year's feedback with direct reports than in years when they did not discuss the prior year's feedback (Walker & Smither, 1999). Hezlett and Ronnkvist (1996) also found that without action planning and feedback to raters, the probability of behavior change is low.

This compares favorably with several undocumented reports that this practice is instrumental in managers effecting behavioral change after they receive feedback (Goldsmith, 2004; Goldsmith & Morgan, 2004; Goldsmith & Underhill, 2001). In addition, Farh et al. (1991) report that asking peers for negative and positive feedback enhances their subsequent evaluations. Finally, Hazucha, Hezlett, and Schneider (1993) have identified three factors that increase the likelihood of behavior change: sharing development plans with

others, getting a positive and encouraging response, and a having a greater variety of developmental experiences.

The organizational context also includes systems and structures that may or may not support behavior change and help maintain it. The presence of a defined developmental goal-setting process is one factor that has been identified as important for making and sustaining changes in leadership behavior (Dalton & Hollenbeck, 2001). London and Smither (1995) reviewed the feedback processes used by twenty organizations and found that only 40% always linked 360-degree feedback to a specific development activity. Of those who did not link developmental activities to their feedback programs, half also did not have a process to ensure that participants set goals based on their feedback (London & Smither, 1995). Regarding maintenance of the desired behaviors, Dalton and Hollenbeck (2001) cite several studies (McCauley & Hughes-James, 1994; Ruderman, Ohlott, & McCauley, 1996) that support the importance of a supportive environment, particularly the supervisor, in maintaining changed behavior. They also state that "… interventions consistently underestimate the amount of organizational and individual resources required for lasting, meaningful change to occur" (Dalton & Hollenbeck, 2001, p. 374). They estimate that five to 10 times the effort and support that is placed on program planning and implementation should be placed on process after the intervention to sustain change. Many of these factors were outside the parameters of this study; however, examining these factors together with personal characteristics of participants and programmatic differences, such as feedback, training, and coaching, might yield a clearer pattern of leadership behavior change in the future.

Implications for Practice. The most significant implication for practice is for training and coaching interventions to include information, discussion, and practice on developmental goal-setting, sharing results and plans with bosses and direct reports, asking for formative feedback, and having a greater variety of developmental experiences. Research shows that these processes maximize the probability of making sustainable behavior change (Farh et al., 1991; Goldsmith, 2004; Goldsmith & Morgan, 2004; Goldsmith & Underhill, 2001; Hazucha et al., 1993; Hezlett & Ronnkvist, 1996; Walker & Smither, 1999). Many feedback-intensive training and development programs currently make such recommendations to participants (Guthrie & Kelly-Radford, 1998; Guthrie & King, 2004; *Intensive Executive Development Workshop*, 2004). It is not clear, however, that these programs provide enough time and skill practice to ensure participants are confident in their abilities and committed to taking these actions. Relatively few studies have collected data on the frequency of these practices or their impact on behavioral outcomes. Data on such environmental factors would facilitate better understanding of the complex process of behavioral change.

Limitations

This study is limited to executive coaching as defined. Findings and recommendations cannot be applied to popular definitions of coaching, such as those employed by the ICF (*International Coach Federation Home Page*, 2004), the ICA (*International Association of Coaches Home Page*, 2003), or CoachVille (*CoachVille Home Page*, 2004). Coaching practices, such as sponsored by these organizations, cover a much broader array of purposes; do not necessarily include any formal assessment of the client's behavior; seldom include a training component; and may be of much shorter duration. These coaches often have varying levels of education, professional certification, or licenses

(*CoachVille Home Page*, 2004; *International Association of Coaches Home Page*, 2003; *International Coach Federation Home Page*, 2004).

The limited sample size used in this study also decreased the generalizability of its findings. Particularly because the feedback and training (B) and feedback, training, and coaching (C) groups had only 21 and 9 participants, respectively; results should be interpreted with caution. Conclusions must be limited to the population for the study: clients of the sponsoring company, a small leadership development consulting firm in the southeast United States. As recommended earlier, findings should be replicated with larger samples and different populations. Due to the small sample size, it was not possible to use statistical methodologies, such as multiple regression, that would have isolated the effects of feedback, training, and coaching.

One reason why the feedback, training, and coaching (C) group was so small was that there were only 31 participants who had completed feedback, training, and executive coaching from 1999 to 2005. Of these, 13 (42%) agreed to participate in the study. It was impossible to compare this agreement rate with the other two groups because marketing representatives and other representatives of client companies actually contacted potential participants. Records were only kept for those who agreed to participate in the feedback only (A) and feedback and training (B) groups. However, pre-tests were available for all executive coaching clients, and it was possible to compare pre-test scores of those who declined or failed to complete the post-test with those who agreed to participate and for whom individual item scores were available. *t* tests for independent samples were conducted on relationship and task factor pre-test means. The results are shown in Table 5.4.

Table 5.4

t *Tests for Participants' and Nonparticipants' Pre-test Scores in Group C*

		Relationship Pre-test Scores				Task Pre-test Scores			
Group	N	Mean	Std Err	T	$p > \underline{t}$	Mean	Std Err	T	$p > \underline{t}$
Partic	9	7.8143623	.7679469			8.4901532	.7015797		
Non-P	20	7.0567362	.9466245	2.10	0.0449	7.8602160	.7691598	2.09	0.0459

Means for both relationship and task pre-test factor scores were significantly different for those who participated in the C group and those who did not participate. For the relationship factor, $t = 2.10$, $p < .05$; for the task factor, $t = 2.09$, $p < .05$. For the relationship pre-test mean score, those who did not participate in the study scored significantly lower than those who did participate. The relationship pre-test mean for those who did participate in the feedback, training, and coaching (C) group was not significantly different from pre-test scores in the feedback only (A) and feedback and training (B) groups.

For the task pre-test mean score, those who did not participate in the study also scored significantly lower than those who did participate. The pre-test mean for those who did participate in the feedback, training, and coaching (C) group was significantly higher than pre-test scores in the feedback only (A) and feedback and training (B) groups (see Table 4.17).

In summary, those who did not participate in the study scored significantly lower on the relationship pre-test factor. Those who did participate in the feedback, training, and coaching group (C) scored significantly higher on the task factor. These differences should be examined further in future research. It may be that lower relationship factor scores cause individuals to be reluctant to participate in a second round of feedback. The high score on the

task factor for the C group occurred in a very small sample and may be unique to this sample. There appear to be some significant differences between those who participated in the study and those who did not. Given the small size of the feedback, training, and coaching (C) group in particular, the findings of this study should be interpreted with caution.

Another source of concern is the attrition between those who completed the pre-test (Group A) or agreed to participate in the study (Groups B and C) but did not complete the post-test. An analysis of those who did not complete the study is shown in Table 5.5.

Table 5.5

Attrition Between Pre-test and Post-test by Group

Group	Original Participants	Number Completed	Attrition %
A	164	71	57
B	70	36	49
C	20	12	40

In the feedback only (A) group, 164 participants agreed to take part in the study, but only 71 completed both the pre-test and post-test. In the feedback and training (B) group, 70 participants agreed to participate in the study. All had completed the pre-test prior to being recruited in the study. A total of 36 actually completed the post-test in this group. In the feedback, training, and coaching (C) group, 20 participants agreed to participate in the study. These participants had also already completed the pre-test before being recruited to participate in the study. A total of 12 completed the pre-test in this group. Totals of those completing the post-test do not agree with totals of those used for statistical tests of significance because a number of participants were discarded, because item totals were not included in pre-tests before February 2001. For group A, 57% failed to complete the post-test. In group B, 49% failed to complete the post-test; and in group C, 40% failed to complete

the post-test. This pattern was expected because participants in the feedback and training (B) and feedback, training, and coaching (C) groups were expected to be more highly motivated to receive a second round of feedback because of participating in training and coaching interventions.

Another issue that suggests caution in interpreting findings was the need to add peer-coaching participants to the executive-coaching group because of such low numbers of participants. Of the 13 participants who originally agreed to participate in the study, only 6 completed the post-test. For this subgroup, attrition was equivalent to 54%, which is similar to the other groups.

As defined in this and other studies, executive coaching included 10 to 20 hours of one-on-one coach interaction with the client over a 6- to 12-month period. Consistent with accepted executive coaching practice, this process consisted of a series of 8 to 12 face-to-face interviews with superiors, peers, and subordinates, which were transcribed and reviewed with each executive coaching participant by a consultant in a one- to two-day face-to-face meeting. The meeting also included developmental goal setting, strategizing, and planning (Kaplan & Palus, 1994). This process followed the participant's attendance at a five-day training program at which he or she received written feedback from colleagues. Additional follow-up meetings to assess progress and adjust strategies and plans were held in person and by phone approximately 6 and 12 months following the coaching intervention.

In the peer coaching group, participants did not receive reports of verbatim interviews of colleagues, but they did participate in a two-day follow-up group meeting with other participants who had attended the original feedback and training intervention. The group meeting included peer observations and peer feedback and gave participants an opportunity

to share information about strengths, weaknesses, and progress on development plans, as well as updating and strategizing new development plans (Baumgartner). Peer observation is defined as being when "… participants observe others throughout … the program [and] … give each other constructive feedback on the impact of their behavior, both during classroom activities and outside of class, during evening activities" (Guthrie & King, 2004, p. 38-39). In peer coaching, participants "… are provided with a model of a new skill set, form groups of three, and are given the opportunity to play coach, learner and observer" (Guthrie & King, 2004, p. 39). The intervention that these two subgroups received may not be equivalent. Peer coaching is a relatively new development in leadership development, and no studies comparing outcomes of the two processes have been performed.

Recommendations for Future Research

Because of the small numbers of participants in the (B) feedback and training (21) and (C) feedback, training, and coaching (9) groups, it is recommended that findings from this study be replicated in larger groups and across different populations. Replicating the differences for relationship factors and task factors would be of great interest, given that exploratory factor analysis of the actual factors present in the responses is currently not a common research practice. Exploring differences in task and relationship behavior changes resulting from feedback, training, or coaching appears to be a potentially interesting new area of inquiry that might facilitate better understanding of this complex process. If differences in task behavior are replicated, the question of why these behaviors change more quickly or more easily than relationship behaviors would also provide the basis for an intriguing investigation.

In order to capture information from all three components—person, program, and environment—that affect behavioral change, a more robust sample would permit analysis by multiple regression techniques. This would be one way to start to understand the complex relationships among all the variables involved. Multiple regression techniques would also allow researchers to isolate the effects of feedback, training, and coaching from each other. In addition, researchers are often counseled to use multiple methods to confirm data (Martineau, 2004). Additional methods, such as interviews with participants or significant others such as bosses, self-report data, psychological tests, behavioral observations, tracking of action plans, observation of team meetings, or climate surveys would provide additional information to confirm pre-test/post-test scores and measure the impact of external components (Martineau, 2004).

Conclusion

While this study did not conclusively prove that feedback, training and or coaching result in leadership behavior change, it did identify several interesting questions for future research. Examining how individual differences in participants' personal characteristics may impact leadership behavior change is one interesting area of inquiry. Environmental characteristics after receiving feedback and training, such as follow-up with one's superior or direct reports, is another interesting question for future research. Further research into how differences in task and relationship behavior affect leadership behavior change would also add to our understanding of this complex area of behavior change. Finally, the question of how participants in executive coaching may differ from those who are offered the opportunity for feedback and or training is one that should be answered by future research.

Footnotes

[1]Over a period of four to 38 months, normal turnover and attrition make it impossible to guarantee that participants were able to distribute their post-test surveys to exactly the same respondents as were surveyed in the pre-test. Because most of the variance in 360-degree feedback ratings is idiosyncratic to the raters (Scullen, Mount, & Goff, 2000), efforts were made to keep the raters constant in both pre- and post-tests by encouraging participants to distribute surveys only to those to whom they distributed surveys in the pre-test. Exceptions were only recommended in order to meet the minimum number of respondents to receive a feedback report (n = 5).

.

References

Agresti, A., & Finlay, B. (1997). *Statistical methods for the social sciences* (Third ed.). Upper Saddle River, NJ: Prentice Hall.

Anderson, M. C. (2001). *Case study on the return on investment of executive coaching*. Raleigh, NC: Pyramid Resource Group.

Atwater, L. A., Roush, P., & Fischthal, A. (1995). The influence of upward feedback on self and follower ratings of leadership. *Personnel Psychology, 48*, 34-59.

Baird, J. A., & Bolton, M. (1999, Mar. 22–23, 1999). *Coaching for breakthroughs.* Paper presented at the Tools/Benchmarks Users Conference, Greensboro, NC.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*, 63-103.

Ballinger, M. S. (2000). Participant self-perceptions about the causes of behavior change from a program of executive coaching. *Dissertation Abstracts International*(61-11A), 4451.

Bass, B. (1985). *Leadership and performance beyond expectations*. New York: Free Press.

Baumgartner, J. M. (personal communication, April 13, 2005).

Bernardin, H. J., Hagan, C., Ross, S., & Kane, J. S. (1995, May 1995). *The effects of a 360-degree appraisal system on managerial performance.* Paper presented at the tenth annual conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Blake, R., & Mouton, J. (1985). *The managerial grid III*. Houston, TX: Gulf Publishing House.

Bolch, M. (2001). Proactive coaching. *Training Magazine, 38*(5), 58-65.

Bracken, D. W., Timmreck, C. W., & Church, A. H. (2001). *Handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes* (First ed.). San Francisco: Jossey-Bass.

Bray, D., Campbell, R., & Grant, D. (1974). *Formative years in business*. New York: John Wiley and Sons, Inc.

Bricklin, S. M. (2001). The RAPPORT program: A model for improving the emotional intelligence of executive coaching clients. *Dissertation Abstracts International, 62-11B*, 5363.

Brown, M. E. (1999). Personality correlates of leadership behaviors. *Dissertation Abstracts International, 60-12B*, 6402.

Brutus, S., Fleenor, J. W., & London, M. (1998). Elements of effective 360-degree feedback systems. In W. W. Tornow, M. London & C. Associates (Eds.), *Maximizing the value of 360-degree feedback: A process for successful individual and organizational development* (pp. 11-27). San Francisco: Jossey-Bass publishers.

Center for Creative Leadership. (2003). Retrieved March 15, 2004, from http://www.ccl.org/CCLCommerce/index.aspx?CatalogID=Home

Chappelow, C. T. (2004). 360-degree feedback. In C. D. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership handbook of leadership development* (Second ed., pp. 58-84). San Francisco: Jossey-Bass.

Charbonneau, M. A. (2002). Media selection in executive coaching: A qualitative study. *Dissertation Abstracts International, 64-01B*, 450.

CoachVille Home Page. (2004). Retrieved March 24, 2004, from http://www.coachville.com/

Collins, D. B., & Holton, E. F., III. (2004). The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001. *Human Resource Development Quarterly, 15*(2), 217-248.

Conger, J. A. (1992). *Learning to lead: The art of transforming managers into leaders*. San Francisco: Jossey-Bass.

Craig, S. B. (2002). *Implicit theories and beta change in longitudinal evaluations of training effectiveness: An investigation using item response theory.* Virginia Polytechnic Institute and State University, Blacksburg, VA.

Craig, S. B., & Kaiser, R. B. (2003). Applying item response theory to multisource performance ratings: What are the consequences of violating the independent observations assumption? *Organizational Research Methods, 6*(1), 44-60.

Crane, T. G. (2001). *The heart of coaching: Using transformational coaching to create a high-performance culture*. San Diego, CA: FTA Press.

Dalton, M. A. (1997). When the purpose of using multi-rater feedback is behavior change. In D. Bracken, M. A. Dalton, R. A. Jako, C. D. McCauley & V. A. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?* (pp. 1-6). Greensboro, NC: Center for Creative Leadership.

Dalton, M. A., & Hollenbeck, G. P. (1996). *How to design an effective system for developing managers and individuals*. Greensboro, NC: Center for Creative Leadership.

Dalton, M. A., & Hollenbeck, G. P. (2001). A model for behavior change. In D. W. Bracken, C. W. Timmreck & A. H. Church (Eds.), *Measuring the impact of multisource feedback: The comprehensive resource for designing and implementing MSF processes*. San Francisco: Jossey-Bass.

Day, D. V. (2000). Leadership development: A review in context. *The Leadership Quarterly, 11*(4), 581-613.

Dore, R. M. (2001). *A comparative study of executive coaching as an alternate adult teaching methodology within a business environment* (ERIC Technical Report). Columbus, OH: Center for Adult, Career and Vocational Education.

Dotlich, D. L., & Cairo, P. C. (1999). *Action coaching: How to leverage individual performance for company success*. San Francisco: Jossey-Bass.

Douglas, C. A. (1997). *Formal mentoring programs in organizations: An annotated bibliography*. Greensboro, NC: Center for Creative Leadership.

Douglas, C. A., & Morley, W. H. (2000). *Executive coaching: An annotated bibliography*. Greensboro, NC: Center for Creative Leadership.

Duffy, E. M. (1984). A feedback-coaching intervention and selected predictors in outplacement. *Dissertation Abstracts International, 45-05B*, 1611.

Dyer, T. J. (2002). Executive development: Outer goals and inner coaching. *Employment Relations Today, 29*(1), 55-61.

Edwards, M. R., & Ewen, A. J. (1996). *360 degree feedback: The powerful new model for employee assessment & performance improvement*. New York: American Management Association.

Fanasheh-Husam, A. (2003). The perception of executive coaching among CEOs of America's top 500 companies. *Dissertation Abstracts International, 64-03A*, 736.

Farh, J. L., Cannella, A. A., & Bedeian, A. G. (1991). Peer ratings, the impact of purpose on rating quality and user acceptance. *Group and Organizational Studies, 16*(4), 367-386.

Farr Associates. (2003). Retrieved March 15, 2004, from http://www.farr.net/services/services.html
Fiedler, F. E. (1971). *Leadership*. New York: General Learning Press.
Filipczak. (1998). The executive coach: Helper or healer? *Training Magazine, 35*(3), 30-36.
Fleenor, J. W., & Prince, J. M. (1997). *Using 360-degree feedback in organizations: An annotated bibliography*. Greensboro, NC: Center for Creative Leadership.
Fournies, F. F. (1987). *Coaching for improved work performance*. Liberty House: Blue Ridge Summit, PA.
Gilley, J. W., & Boughton, N. W. (1996). Stop managing, start coaching! How performance coaching can enhance commitment and improve productivity. Chicago: Irwin.
Goldsmith, M. (2000). Coaching for behavioral change. In M. Goldsmith, L. Lyons & A. Freas (Eds.), *Coaching for leadership: How the world's greatest coaches help leaders learn* (pp. 21-25). San Francisco: Jossey-Bass/Pfeiffer.
Goldsmith, M. (2004). *Impact of direct report feedback*. Retrieved Feb 28, 2004, from http://marshallgoldsmith.com/html/articles/impact.html
Goldsmith, M., Lyons, L., & Freas, A. (2000). *Coaching for leadership: How the world's greatest coaches help leaders learn*. San Francisco: Jossey-Bass/Pfeiffer.
Goldsmith, M., & Morgan, H. (2004). *Increasing leadership effectiveness*. Retrieved Feb 28, 2004, from http://marshallgoldsmith.com/html/articles/increasing.html
Goldsmith, M., & Underhill, B. O. (2001). Multisource feedback for executive development. In D. W. Bracken, C. W. Timmreck & A. H. Church (Eds.), *Handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes* (First ed.). San Francisco: Jossey-Bass.
Guthrie, V. A., & Kelly-Radford, L. (1998). Feedback-intensive programs. In E. Van Velsor, C. D. McCauley & R. S. Moxley (Eds.), *The Center for Creative Leadership handbook of leadership development* (pp. 66-105). San Francisco: Jossey-Bass.
Guthrie, V. A., & King, S. N. (2004). Feedback-intensive programs. In C. D. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership Handbook of Leadership Development. Second Edition* (pp. 25-57). San Francisco: Jossey-Bass.
Hagberg, R. (1996). Identify and help executives in trouble. *HR Magazine, 41*(8), 88-93.
Hall, D. T., Otazo, K. I., & Hollenbeck, G. P. (1999). Behind closed doors: What really happens in executive coaching. *Organizational Dynamics, 27*(3), 39-53.
Hamilton, K. (1996). Need a life? Get a coach. *Newsweek, 127*(6), 48.
Hatcher, L., & Stepanski, E. J. (1994a). *A step-by-step approach to using SAS for factor analysis and structural equation modeling*. Cary, NC: SAS Institute Inc.
Hatcher, L., & Stepanski, E. J. (1994b). *A step-by-step approach to using the SAS system for univariate and multivariate statistics*. Cary, NC: SAS Institute Inc.
Hazucha, J. F., Hezlett, S. A., & Schneider, R. J. (1993). The impact of 360-degree feedback on management skills development. *Human Resource Management, 32*, 325-352.
Hazucha, J. F., Symanski, C., & Birkeland, S. (1992). *Will my boss see my ratings? Effect of confidentiality on self-boss congruence.* Paper presented at the Symposium of the American Psychological Association, Washington, DC.
Hegarty, W. H. (1974). Using subordinate ratings to elicit behavioral changes in managers. *Journal of Applied Psychology, 59*, 764-766.

Hellervik, L. W., Hazucha, J. F., & Schneider, R. J. (1992). Behavior change: Models, methods, and a review of evidence. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3). Palo Alto, CA: Consulting Psychologists Press.

Hersey, P., & Blanchard, K. (1988). *Management of organizational behavior*. New York: Prentice Hall.

Hezlett, S. A., & Ronnkvist, A. M. (1996, April). *The effects of multi-rater feedback on managers; skill development: Factors influencing behavior change.* Paper presented at the Society for Industrial and Organizational Psychology, San Diego, CA.

Hudson, F. M. (1999). *The handbook of coaching: A comprehensive resource guide for managers, executives, consultants, and human resource professionals*. San Francisco: Jossey-Bass Publishers.

Hughes, J. L. (2002). Adjusting the mirror: Strategies for coaching executives with narcissistic personality features. *Dissertation Abstracts International, 63-08B*, 3966.

Intensive Executive Development Workshop. (2004). High Point, NC: Farr Associates.

International Association of Coaches Home Page. (2003). Retrieved March 24, 2004, from http://www.certifiedcoach.org/

International Coach Federation Home Page. (2004). Retrieved March 24, 2004, from http://www.coachfederation.org/

Johnson, J. W., & Ferstl, K. L. (1999). The effects of interrater and self-other agreement on performance improvement following upward feedback. *Personnel Psychology, 52*, 271-303.

Johnson, T., & Berk, K. N. (2000). *Data analysis with JMP IN 4*. Pacific Grove, CA: Duxbury Thomson Learning.

Jossi, F. (1999). On the side lines. *Human Resource Executive, 13*(1), 72-75.

Kampa-Kokesch, S. (2001). Executive coaching as an individually tailored consultation intervention: Does it increase leadership? *Dissertation Abstracts International, 62-07B*, 3408.

Kaplan, R. E., Drath, W. H., & Kofodimos, J. R. (1987). High hurdles: The challenge of executive self-development. *Academy of Management Executive, 1*, 195-205.

Kaplan, R. E., Drath, W. H., & Kofodimos, J. R. (1991). *Beyond ambition: How driven managers can lead better and live better*. San Francisco: Jossey-Bass.

Kaplan, R. E., & Palus, C. J. (1994). *Enhancing 360-degree feedback for senior executives*. Greensboro, NC: Center for Creative Leadership.

Kilburg, R. R. (1996). Toward a conceptual understanding and definition of executive coaching. *Consulting Psychology Journal: Practice and Research, 48*(2), 134-144.

Kilburg, R. R. (2000). *Executive coaching: Developing managerial wisdom in a world of chaos*. Washington, D. C.: American Psychological Association.

Kluger, A. N., & DeNisi, A. (1966). The effects of feedback interventions on performance: A historical review, meta-analysis, and a preliminary feedback theory. *Psychological Bulletin, 119*, 254-284.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, meta-analysis, and a preliminary feedback theory. *Psychological Bulletin, 119*, 254-284.

Knodt, G. J. (1990). *Effects of feedback and coaching on the transfer of people management skills from the cognitive to the behavioral state.* Columbia Pacific University, Novato, CA.

Leonard, T. J. (1999). *Becoming a coach: The Coach U approach*. Steamboat Springs, CO: Coach U Press.

Leslie, J. B., & Fleenor, J. W. (1998). *Feedback to managers: A review and comparison of multi-rater instruments for management development* (Third ed.). Greensboro, NC: Center for Creative Leadership.

Leslie, J. B., & Van Velsor, E. (1996). *A look at derailment today: North America and Europe*. Greensboro, NC: Center for Creative Leadership.

Littell, R. C., Fruend, R. J., & Spector, P. C. (1991). *SAS system for linear models* (Third edition ed.). Cary, NC: SAS Institute Inc.

Lombardo, M. M., & Eichinger, R. W. (1989). *Preventing derailment: What to do before it's too late*. Greensboro, NC: Center for Creative Leadership.

Lombardo, M. M., McCauley, C. D., by, R., Moxley, R. S., Dalton, M. A., & Wacholz, J. (1994). *Benchmarks: Developmental reference points. A manual and trainer's guide*. Greensboro, NC: Center for Creative leadership.

London, M. (2001). The great debate: Should multisource feedback be used for administration or for development only? In M. London (Ed.), *How people evaluate others in organizations*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

London, M., & Smither, J. W. (1995). Can multisource feedback change self-evaluation, skill development, and performance? Theory-based applications and directions for research. *Personnel Psychology, 48*, 803-839.

London, M., Smither, J. W., & Adsit, D. J. (1997). Accountability: The Achilles' heel of multisource feedback. *Group and Organization Management, 22*(2), 162-184.

London, M., & Tornow, W. W. (1998). 360-degree feedback--more than a tool! In W. W. Tornow, M. London & C. Associates (Eds.), *Maximizing the value of 360-degree feedback: A process for successful individual and organizational support* (pp. 1-8). San Francisco: Jossey-Bass Publishers.

Luthans, F., & Peterson, S. J. (2003). 360-degree feedback with systematic coaching: Empirical analysis suggests a winning combination. *Human Resource Management, 42*(3), 243-256.

Martineau, J. W. (1998). Using 360-degree surveys to assess change. In W. W. Tornow, M. London & C. Associates (Eds.), *Maximizing the value of 360-degree feedback: A process for successful individual and organizational development* (pp. 217-248). San Francisco: Jossey-Bass Publishers.

Martineau, J. W. (2004). Evaluating the impact of leader development. In C. D. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership handbook for leadership development* (Second ed.). San Francisco: Jossey-Bass.

McCall, M. W., Jr., & Lombardo, M. M. (1983). *Off the track: Why and how successful executives get derailed* (Technical report no. 21). Greensboro, NC: Center for Creative Leadership.

McCall, M. W., Jr., Lombardo, M. M., & Morrison, A. M. (1988). *The lessons of experience: How successful individuals develop on the job*. San Francisco: Jossey-Bass.

McCauley, C. D., & Hughes-James, M. (1994). *An evaluation of outcomes of a leadership development program*. Greensboro, NC: Center for Creative Leadership.

McCauley, C. D., & Van Velsor, E. (2004). *The Center for Creative Leadership handbook of leadership development* (Second ed.). San Francisco: Jossey-Bass.
Meyer, H. H., Kaye, E., & French, J. R. P. J. (1965). Split roles in performance appraisal. *Harvard Business Review*(65108).
Miller, D. J. (1989). The effect of managerial coaching on transfer of training. *Dissertation Abstracts International, 50-08A*, 2435.
Morris, B. (2000). So you're a player. Do you need a coach? *Fortune, 141*(4), 144-154.
Nowack, K. M., & Wimer, S. (1997). Coaching for human performance. *Training & Development, 51*(10), 28-32.
Olivero, G., Bane, K. D., & Kopelman, R. E. (1997). Executive coaching as a transfer of training tool: Effects on productivity in a public agency. *Public Personnel Management, 26*(4), 461-470.
Orenstein, R. L. (2000). Executive coaching: an integrative model. *Dissertation Abstracts International, 61-04B*, 2257.
Patton, C. (2001a). Cyber coach. *Human Resource Executive, 15*(3), 52-58.
Patton, C. (2001b). Rating the returns. *Human Resource Executive, 15*(5), 40-43.
Peltier, B. (2001). *The psychology of executive coaching: Theory and application*. New York: Brunner-Routledge.
Peterson, D. B. (1993). Skills learning and behavior change in an individually tailored management coaching and training program. *Dissertation Abstracts International, 54-03B*, 1701.
Peterson, D. B., & Hicks, M. D. (1996). *Leaders as coaches: Strategies for coaching and developing others*. Minneapolis: Personnel Dimensions International.
Peterson, D. B., & Hicks, M. D. (1999). Strategic coaching: Five ways to get the most value. *HR Focus, 76*(2), S7-S8.
Phillips, J. J. (2003). *Return on investment in training and performance improvement programs. Second edition*. Woburn, MA: Butterworth-Heinemann.
Pulley, M. L., Sessa, V. I., & Malloy, M. (2002). E-leadership: A two-pronged idea. *Training and Development, 56*(3), 34-47.
Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). *Applied regression analysis: A research tool* (Second edition ed.). New York: Springer-Verlag New York, Inc.
Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. (1996). A longitudinal study of upward feedback. *Personnel Psychology, 49*, 599-612.
Ruderman, M. N., Ohlott, P. J., & McCauley, C. D. (1996, Apr.). *Developing from job experiences: The role of self-esteem and self-efficacy.* Paper presented at the Society for Industrial and Organizational Psychology, San Diego, CA.
Sawczuk, M. P. (1990). Transfer-of-training: Reported perceptions of participants in a coaching study in six organizations (management development). *Dissertation Abstracts International, 51-12A*, 4195.
Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology, 85*, 956-970.
Smither, J. W., London, M., Flautt, R., Vargas, Y., & Kuchine, I. (2003). Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. *Personnel Psychology, 56*, 23-43.

Smither, J. W., & Reilly, S. P. (2001). Coaching in organizations: A social psychological perspective. In M. London (Ed.), *How people evaluate others in organizations*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Smither, J. W., & Walker, A. G. (1995). An examination of the effects of an upward feedback program over time. *Personnel Psychology, 48*, 1-34.

Smither, J. W., & Walker, A. G. (2001). Measuring the impact of multisource feedback: The comprehensive resource for designing and implementing MSF processes. In D. W. Bracken, C. W. Timmreck & A. H. Church (Eds.), *Handbook of multisource feedback* (First ed.). San Francisco: Jossey-Bass.

Stowell, S. J., & Starcevich, M. M. (1998). *The coach: Creating partnerships for a competitive edge*. Sandy, UT: The Center for Management and Organization Effectiveness.

Sztucinski, K. (2001). The nature of executive coaching: An exploration of the executive's experience. *Dissertation Abstracts International, 62-10B*, 4826.

Thach, L. (2002). The impact of executive coaching and 360-degree feedback on leadership effectiveness. *Leadership & Organization Development Journal, 23*(4), 205-214.

Thach, L., & Heinselman, T. (1999). Executive coaching defined. *Training & Development, 53*(3), 34-29.

Thompson, A. D., Jr. (1986). A formative evaluation of an individualized coaching program for business managers and professionals. *Dissertation Abstracts International, 47-12A*, 4339.

Ting, S., & Hart, E. W. (2004). Formal coaching. In C. D. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership handbook of leadership development* (Second edition ed., pp. 116-150). San Francisco: Jossey-Bass.

Tornow, W. W., London, M., & Associates, C. (1998). *Maximizing the value of 360-degree feedback: A process for successful individual and organizational development*. San Francisco: Jossey-Bass.

Turner, C. E. (2003). Executive coaching as a leadership development strategy. *Dissertation Abstracts International, 64004A*, 1332.

Van Velsor, E., Moxley, R. S., & Bunker, K. A. (2004). The leader development process. In C. A. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership handbook of leadership development* (Second ed.). San Francisco: Jossey-Bass.

Vicere, A. A., & Fulmer, R. M. (1997). *Leadership by design*. Boston: Harvard Business School Press.

Walker, A. G., & Smither, J. W. (1999). A five-year study of upward feedback: What managers do with their results matters. *Personnel Psychology, 52*, 393–423.

Witherspoon, R., & White, R. P. (1997). *Four essential ways that coaching can help executives*. Greensboro, NC: Center for Creative Leadership.

Witherspoon, R., White, R. P., & Hutchinson, P. G. (1996). Coaching as a collaboration. *Training & Development, 50*(3), 14-16.

Yammarino, F. J. (2003). Modern data analytic techniques for multisource feedback. *Organizational Research Methods, 6*(1), 6-14.

Young, D. P., & Dixon, N. M. (1996). *Helping leaders take effective action: A program evaluation*. Greensboro, NC: Center for Creative Leadership.

APPENDICES

Appendix A

DEMOGRAPHIC DATA FORM for Participants Record #: _____

Information will be kept confidential. Demographic data is for research purposes only. We ask for your name on this form so that we can correlate this research information with your personality test scores and 360 results. When this information is entered in the research database, your name will be purged and only a research record number assigned to ensure confidentiality.

Name _____ Date _____

Age _____ ❒single ❒ married ❒ divorced ❒ widowed ❒separated

Gender: ❒Male ❒Female

Please circle highest education completed: → High School 1 2 3 4 College 1 2 3 4 5 6 7 8 9+

Years in Present Job: ____ Years in Present Company:____ Years in Present Industry: _____

Please check the number of employees in your organization:

- ❒ Fewer than 50 employees
- ❒ 51-250 employees
- ❒ 251-500 employees
- ❒ 501-1000 employees
- ❒ More than 1001 employees

Please check your industry:

- ❒ Finance, Insurance and Real Estate Institutions
- ❒ Wholesale Trades
- ❒ Retail Trade (building materials, food stores, restaurants, etc.)
- ❒ Manufacturing
- ❒ Agriculture, Forestry, and Fishing
- ❒ Mining
- ❒ Construction
- ❒ Services (hotels, health services, legal services, accounting)

Please check your level in the organization:

- ❒ Top Tier: Chief Level Officer/President
- ❒ Executive Tier: SVP, EVP, Directors
- ❒ Upper-Middle Tier: Department Head, Plant Manager, etc.
- ❒ Middle Management Tier: Supervisor, Prof. Staff, Team Leader
- ❒ Non-Management Tier

Please give us some information about how you used the 360-degree feedback you have received. This will help us identify what techniques are most helpful. The information you share will be kept in strictest confidence. It will only be reported as group averages and will not be connected with your name or any identifying information.

After you received your feedback, did you set any goals for any behaviors you wanted to change? ❒No ❒Yes

If you set goals, approximately how many did you set? 0 1 2 3 4 5 or more

How specific do you feel your goals were?

❒ Very General ❒ Somewhat General ❒ Not General, or Specific ❒ Somewhat Specific ❒ Very Specific

Please turn the page for some additional questions.

Please check the types of goals for changing your behaviors that you set. Just check the ones you can remember.

- ❒ Communication
  - __ Listening and questioning
  - __ Articulating information
- ❒ Interpersonal Skills
  - __ Sensitivity to others
  - __ Shows tact
  - __ Wins cooperation
  - __ Earns trust and loyalty
- ❒ Teamwork
  - __ Helps and provides resources
  - __ Acts as member of team
  - __ Puts team before self-interest
  - __ Gives credit to others
  - __ Avoids politics
- ❒ Initiative
  - __ Seeks new challenges
  - __ Self-starter and finisher
  - __ Overcomes obstacles
- ❒ Creativity
- ❒ Adaptability
  - __ Open to new ideas
  - __ Flexible with differences
  - __ Deals well with uncertainty
- ❒ Judgment
  - __ Consults others
  - __ Considers alternatives and consequences
  - __ Makes tough decisions
  - __ Does not procrastinate
  - __ Is ethical
- ❒ Leadership
  - __ Provides purpose, vision, values
  - __ Is responsible for own actions
  - __ Inspires others to greater performance
- ❒ Maturity
  - __ Deals well with stress and frustration
  - __ Projects poise and confidence

Did you meet with your boss to discuss your feedback? ❒Yes ❒No

Did your boss assist you in planning how to achieve your goals? ❒Yes ❒No

Did you ask your boss for any support to help you achieve your goals? ❒Yes ❒No

Did your boss provide any additional support to help your achieve your goals? ❒Yes ❒No

Did you meet with your direct reports to discuss your feedback? ❒Yes ❒No

Did your direct reports assist you in planning how to achieve your goals? ❒Yes ❒No

Did you ask your direct reports for any support to help you achieve your goals? ❒Yes ❒No

Did your direct reports provide any additional support to help your achieve your goals? ❒Yes ❒No

Do you feel you have become more effective (or less effective) as a leader since you first received feedback? (Please rate yourself on effectiveness concerning aspects of leadership you can control. Please do not consider environmental or organizational factors that are beyond your control.)

Less Effective					No Perceptible Change					More Effective
-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5

If you want to, please list indicators of your increased leadership effectiveness that you have noticed or measured.

Thank you for participating in this study. The information you have provided will help us learn what techniques are most helpful in enhancing leadership effectiveness. Please be assured that all information will be held in strictest confidence.

RETURN TO: Research Study

Appendix B. Feedback Instrument

Research Study

Leadership Feedback for FULL NAME

NAME *will be participating in a leadership development study with XXX and North Carolina State University. The purpose of the study is to compare the effects of feedback, training, and coaching on leadership behavior change. We are asking for your open and honest evaluation in the dimensions listed below. Your evaluation will be anonymous, and completing this evaluation is entirely voluntary. It will have no effect on your employment.*

The items below are not intended to form the basis for an in-depth leadership effectiveness analysis, nor are they designed to provide feedback in terms of specific behaviors and patterns that need to be examined. They are instead intended to provide a benchmark for self-assessment and planning to enhance leadership effectiveness.

For each item below **choose** a number from **0** (**lowest** effectiveness level) to **10** (**highest** effectiveness level) indicating how well you think she/he does in this dimension, and write or type it in the corresponding box to the right. *Don't spend a lot of time on each statement; just indicate what comes to mind as you read it.*

When you've completed this form, please send it and one copy of the completed informed consent form to the person who has been designated to collect these forms anonymously to send them to Farr. ***Please don't send or copy it to anyone in your organization other than that designated person.*** Thank you!

Communication:	
Demonstrates effective listening and questioning skills	
Effectively articulates information	
Interpersonal Skills:	
Sensitive to needs of others	
Shows tact in dealing with others	
Successfully wins cooperation of others	
Earns trust and loyalty	
Teamwork:	
Willingly helps and provides resources to others	
Acts as a member of the team	
Places the team before personal interest	
Gives credit to others	
Avoids "playing politics"	
Initiative:	
Seeks new challenges and responsibilities	
Self-starter and -finisher	
Persistent in overcoming obstacles	
Creativity:	
Develops innovative ideas and approaches that work	
Adaptability:	
Open to new ideas and approaches	
Flexible in dealing with different viewpoints and styles	
Deals well with uncertainty and ambiguity	
Judgment:	
Consults with others in dealing with difficult situations	
Considers alternatives/consequences before making decisions	
Willing/able to make tough decisions when required	
Does not procrastinate	
Understands and practices appropriate ethics	
Leadership:	
Provides purpose, values, vision	
Takes responsibility for his/her actions	
Earns trust and loyalty; inspires others to greater performance	
Maturity:	
Effectively deals with stress and frustration	
Projects poise and confidence	

© Copyright 2003 by XXX; All Rights Reserved.

Appendix C. Pre-test Letter

Instructions for Research Study (Group A1)

Thank you for agreeing to participate in our research study. As a part of the study, you will receive two feedback data summaries, approximately six months apart. You can compare these two data summaries to see what changes in your leadership effectiveness you have made. Because this is a research study, we will be asking you to complete an informed consent form, in addition to completing procedures for collecting your feedback.

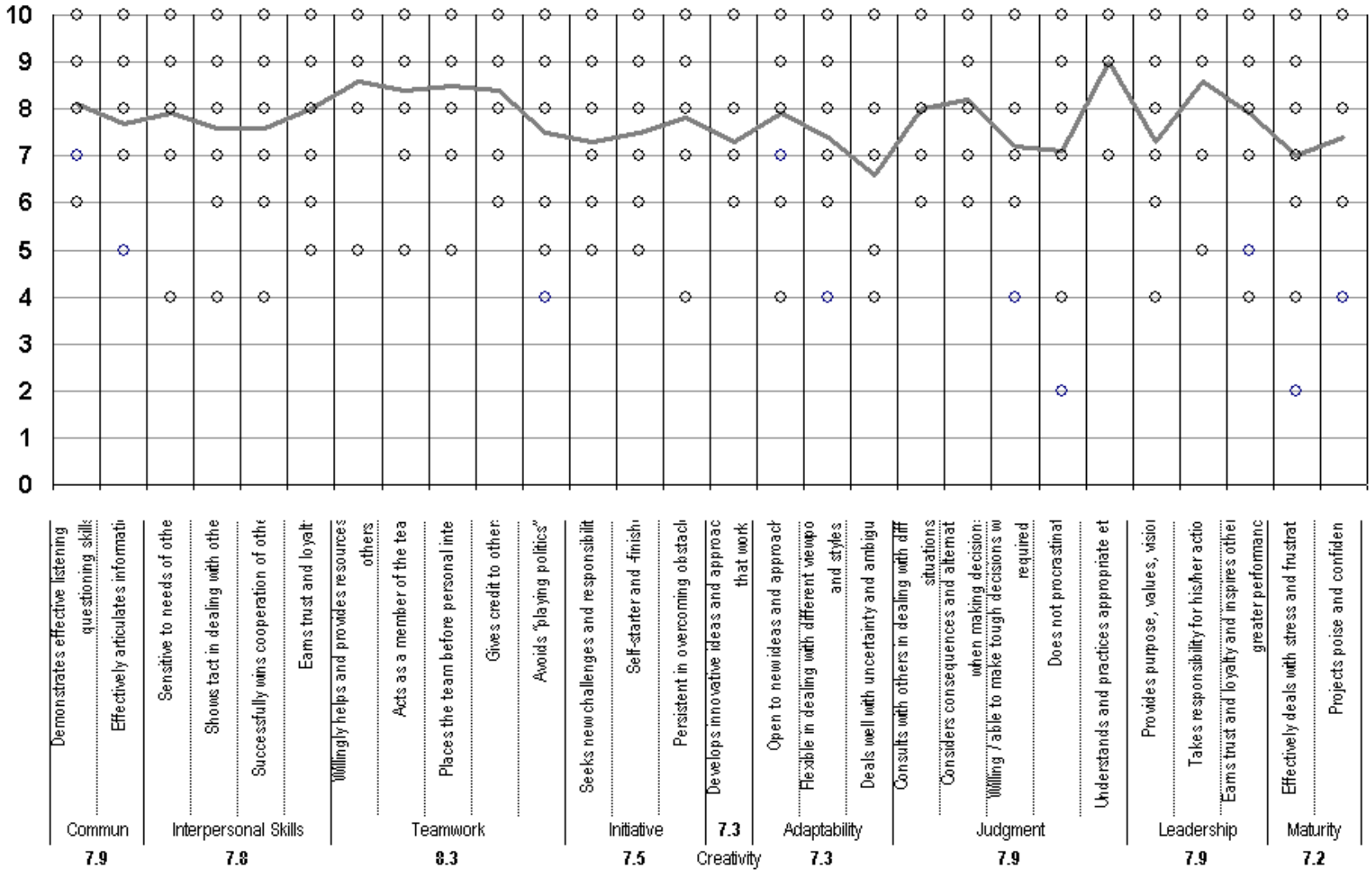
We'd like you to help gather some data on how your leadership impact is perceived by those you work with. It will be compiled by XXX and provided to you in a feedback data summary. The data are confidential between you and XXX and will not be shared with anyone in your organization; we have set up the data gathering process to help keep the information anonymous. There are six simple steps to complete the process on the checklist below.

- ❒ **1** Please review the attached informed consent form and sign two copies. Keep one for your records and send the other to XXX in the small preaddressed envelope.
- ❒ **2** Make a list of individuals (superiors, subordinates, and/or peers) who are in a position to provide valid observations on your leadership impact—people who will complete the form openly and honestly.
- ❒ **3** Select a trusted individual in your organization to collect and return the completed forms, someone who is willing to do this but who will not reveal to **anyone** (including yourself) the names of those who do or don't turn in the completed forms.
- ❒ **4** Fill in your own name at the top of the Leadership Impact Feedback (LIF) form and fill in the name of the individual you have selected to collect your completed LIF forms in the space provided at the bottom of the LIF. Run enough copies (both sides) for the number of people you select: minimum 8, maximum 15.
- ❒ **5** Give your list to this individual. Explain that he/she should get a completed form from each person listed, and ask him/her to be "proactive" in rounding up all the forms **within two weeks**, but **not to tell you who has or hasn't handed one in.** Ask her/him to mail the completed forms in the large preaddressed envelope and then destroy the list of names.
- ❒ **6** Give a feedback form along with a plain envelope to each of the people on your list, personally if possible. Ask them to complete the assessment, put it into the envelope, seal the envelope, and give the **sealed** envelope to the person you have designated to collect the forms. Please ask them to complete this within **one** week. Tell them how much you are counting on their being open and objective in their assessment of your leadership impact.

The staff at XXX will compile the data and provide you with a confidential feedback data summary. The data will also be used in a research study to examine the effects of feedback, training and coaching on leadership behavior change. Participants in the research study will not be identified in any way.

Thanks for your help. The findings of the research study will help us identify techniques that help leaders increase their leadership effectiveness. We hope your new feedback report will help you assess your leadership effectiveness and make changes that enhance your effectiveness.

Appendix D. Sample Feedback Report



10 forms received from the people whom you asked to participate in this assessment have been analyzed and compiled to produce the chart above. Ratings received (**10** highest, **0** lowest) are represented by open dots (**any dot may represent a rating given by more than one person**). The average of all ratings is represented by the gray line. Average ratings for each leadership category are indicated as a bold-faced number at the bottom of that category's columns.

Appendix E. Feedback Summary (Group A1)

Thank you for agreeing to participate in our research study. Enclosed is your first feedback data summary. It shows you the average responses for all individuals who completed Leadership Impact Feedback forms for you. If this is your first opportunity to receive feedback, here are some suggestions you may wish to consider as you review your feedback.

- Most feedback will contain some surprises. Strengths that you were unaware of are pleasant surprises. Your feedback may also contain some blind spots, areas for improvement that you were not aware of. The value of feedback is twofold: to affirm our strengths, so that we continue to use these successful practices; and to point out areas for improvement that may have been overlooked.
- You may have some strong feelings about feedback that comes as a surprise. You may be shocked or even angry about some of your feedback. You may want to reject it or rationalize it, if you do, you will negate the value of the data and of accepting the reality that people may see you differently from how you thought they did.
- After you review the feedback for general patterns, identify at least two strengths that your feedback identifies. Think about what you do to cause people to react positively. For example, if "gives credit to others" is one of your strengths, think about how you acknowledge others' efforts. Review how you do this and create ways to leverage your strengths even more, perhaps by teaching them to others, putting yourself in positions that highlight your key strengths, or utilizing them to address areas for development.
- Then try to focus on a short list of two or three areas in which you would like to improve your performance. It may be tempting to try to address more areas, but research shows that people who concentrate on just a few areas are more successful in improving their performance and overall perceptions of their leadership.
- Set goals for improvement for the areas you have targeted. As with any goal, try to be specific and take a few minutes to identify some strategies you can use. For instance, if you are working on your listening skills, you might develop a strategy to make sure you paraphrase co-workers' suggestions before you discuss how to implement them. And set some targets and identify some indicators you can measure in order to track your progress toward your goals.
- Thank the people who provided feedback. Let them know you are planning to make some changes based on their feedback. You can share your main goals with them, but don't overload them with more information than they want. Remember, it is your feedback, and you have more interest in the details.
- Check in with your feedback providers and ask them to give you feedback on how you are doing regularly. You can do this one on one or at regular staff meetings.
- Build some rewards in for yourself and celebrate your successes.

Thanks for your help. The findings of the research study will help us identify techniques that help leaders increase their leadership effectiveness. We hope your new feedback report will help you assess your leadership effectiveness and make changes that enhance your effectiveness.

In about six months, we will send you another feedback packet, so you can repeat the feedback process. You will again distribute LIF forms to your colleagues. Try to give them to the same people who completed feedback for you the first time. I will mail you a report comparing your feedback summaries so you can see the impact of changes you have made.

In the meantime, if you have any questions or would like further support regarding your feedback results, please contact the researcher, Nona Saling, at (919) 419-7888 or at nsaling@mindspring.com.

Appendix F. Post-test Letter (Groups B, C)

Instructions for Research Study

Thank you for agreeing to participate in our research study. As a part of the study, you will receive a second feedback data summary that you can compare to the feedback data you received when you attended the leadership development workshop at XXX to see what changes in your leadership effectiveness you have made. Because this is a research study, we also need you to complete the attached informed consent form. In addition, we need to ask you to provide some demographic information and to answer a few questions about how you used your initial feedback.

Please follow the same procedure to collect your feedback data that you used for your original feedback data. The data are confidential between you and XXX and will not be shared with anyone in your organization. We have set up the data-gathering process to help keep the information anonymous. There are seven simple steps in the checklist below:

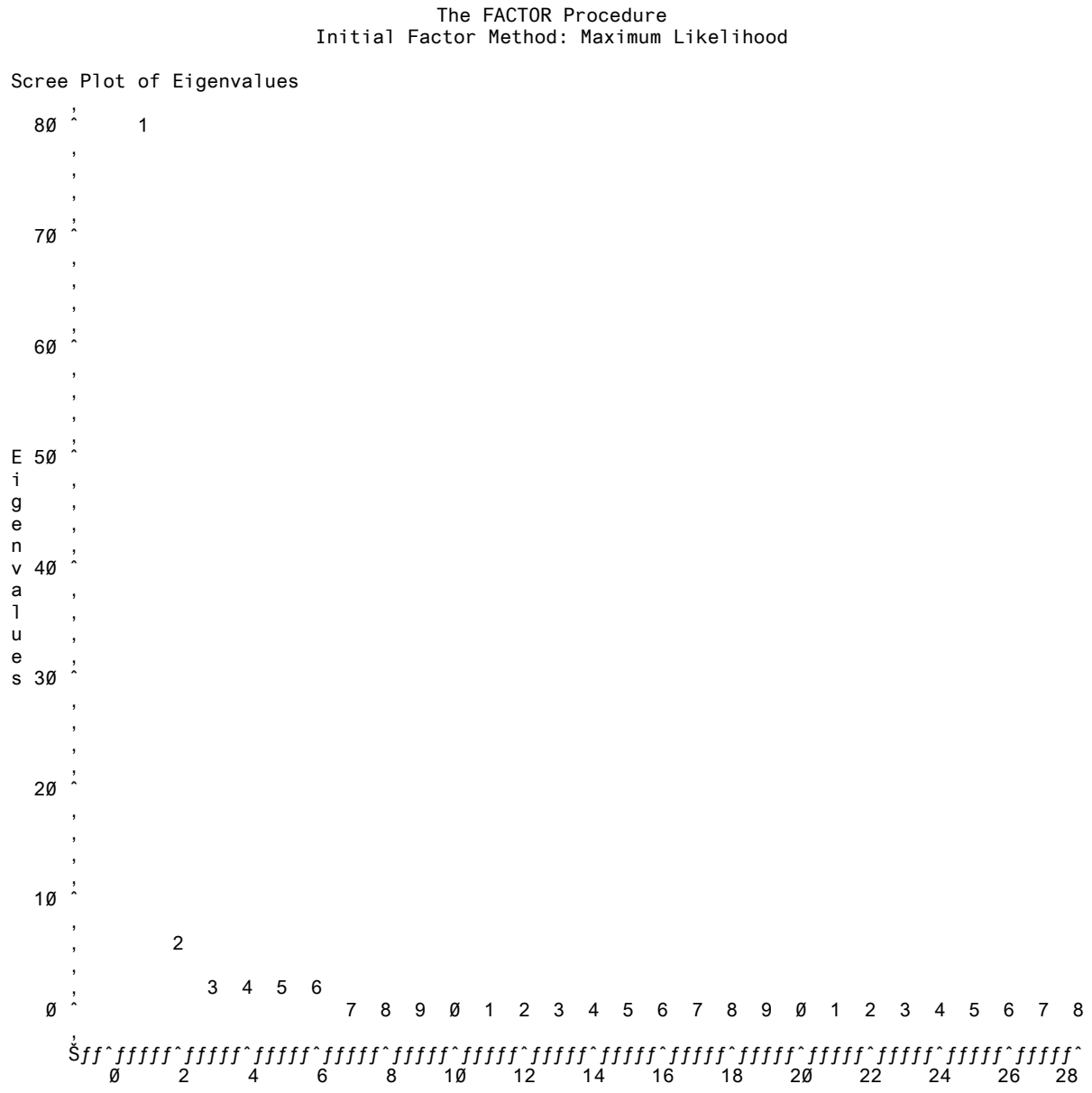
- ❒ **1** Please review the attached informed consent form and sign two copies. Keep one for your records and send the other to XXX in the attached small preaddressed envelope.
- ❒ **2** Please complete both sides of the attached Demographic Data Form for Participants and return it to XXX in the attached small preaddressed envelope.
- ❒ **3** Select a trusted individual in your organization to collect and return the completed forms, someone who is willing to do this but who will not reveal to **anyone** (including yourself) the names of those who do or don't turn in the completed forms.
- ❒ **4** Fill in your own name at the top of the Leadership Impact Feedback (LIF) form, and fill in the name of the individual you have selected to collect your completed LIF forms in the space provided at the bottom of the LIF. Run enough copies (both sides) for the number of people you select: minimum 8, maximum 15.
- ❒ **5** Select a trusted individual in your organization to collect and return the completed forms, someone who is willing to do this but who will not reveal to **anyone** (including yourself) the names of those who do or don't turn in the completed forms.
- ❒ **6** Give your list to this individual. Explain that he/she should get a completed form from each person listed, and ask him/her to be "proactive" in rounding up all the forms **within two weeks**, but **not to tell you who has or hasn't handed one in.** Ask her/him to mail the completed forms in the large preaddressed envelope and then destroy the list of names.
- ❒ **7** Give a feedback form and a plain envelope to each of the people on your list, personally if possible. Ask them to complete the assessment, put it into the envelope, seal the envelope, and give the **sealed** envelope to the person you have designated to collect the forms. Please ask them to complete this within **one** week. Tell them how much you are counting on their being open and objective in their assessment of your leadership impact.

The staff at XXX will compile the data on the forms received and provide you with a confidential feedback data summary. The data will also be used in a research study to examine the effects of feedback, training, and coaching on leadership behavior change. Participants in the research study will not be identified in any way in the research study.

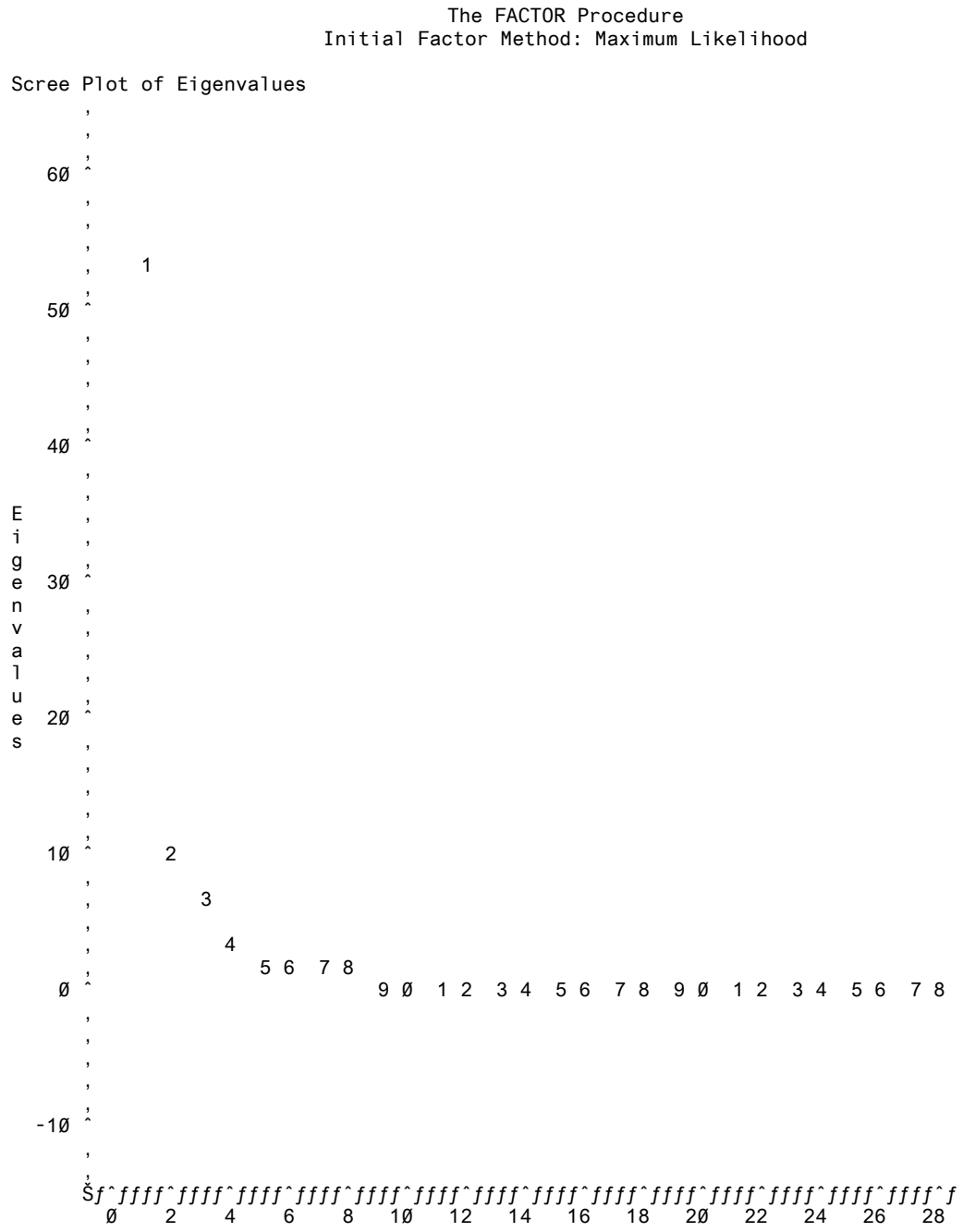
Thanks for your help. The findings of the research study will help us identify techniques that help leaders increase their leadership effectiveness. We hope your new feedback data will help you assess the changes you have made in your leadership effectiveness and continue to enhance your effectiveness.

Appendix G Scree Charts

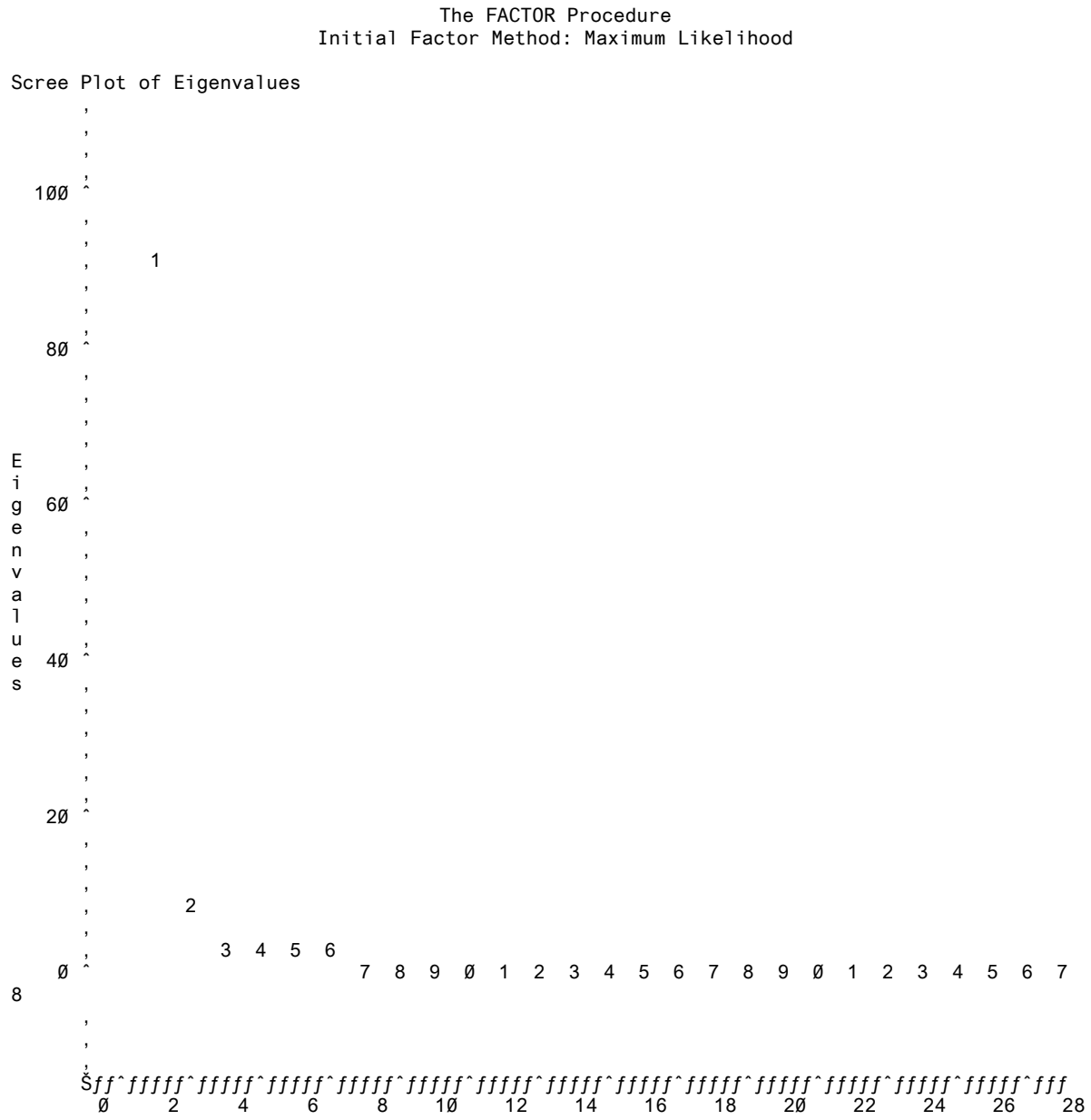
Group A Pre-test Responses



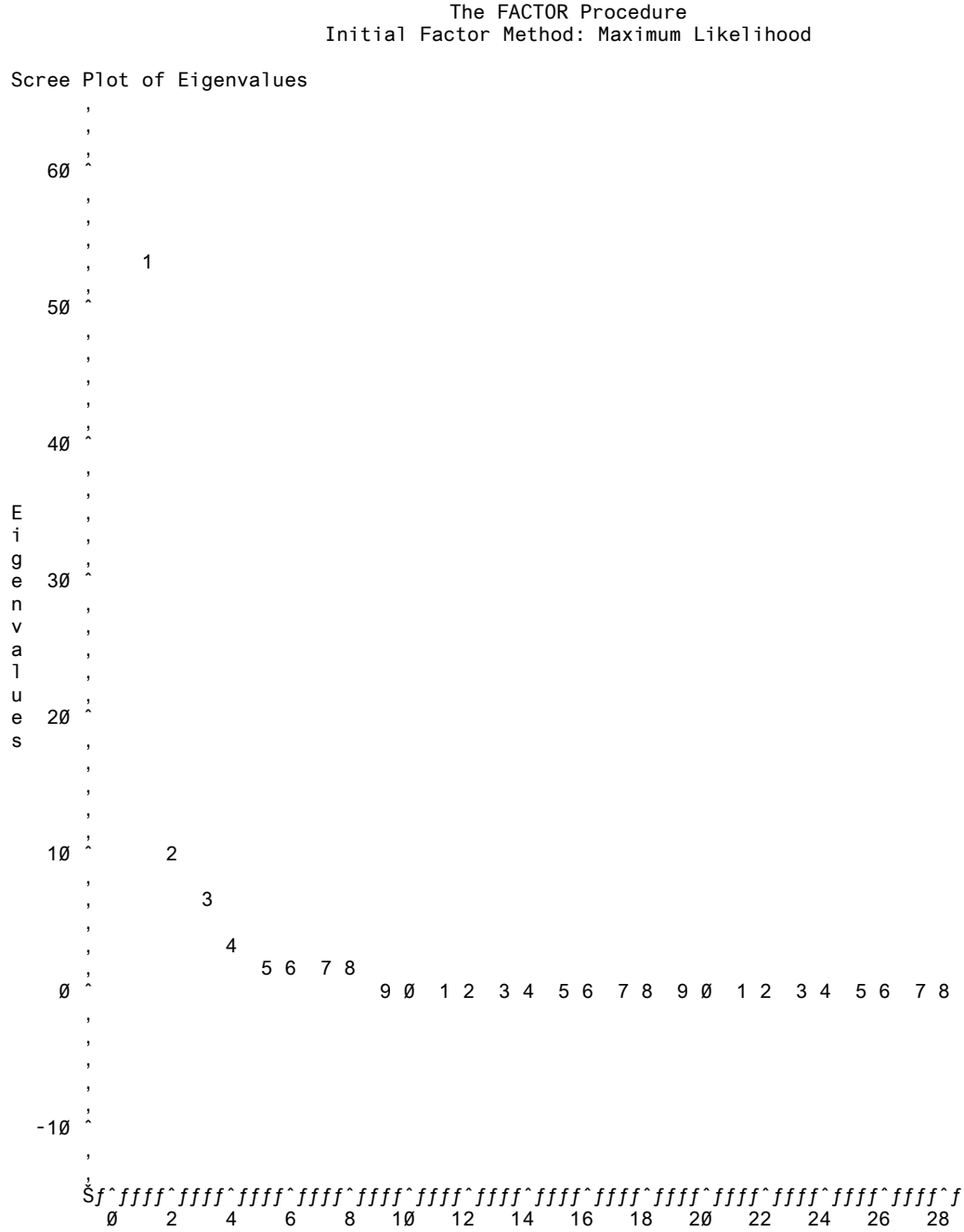
Group A Post-test Responses



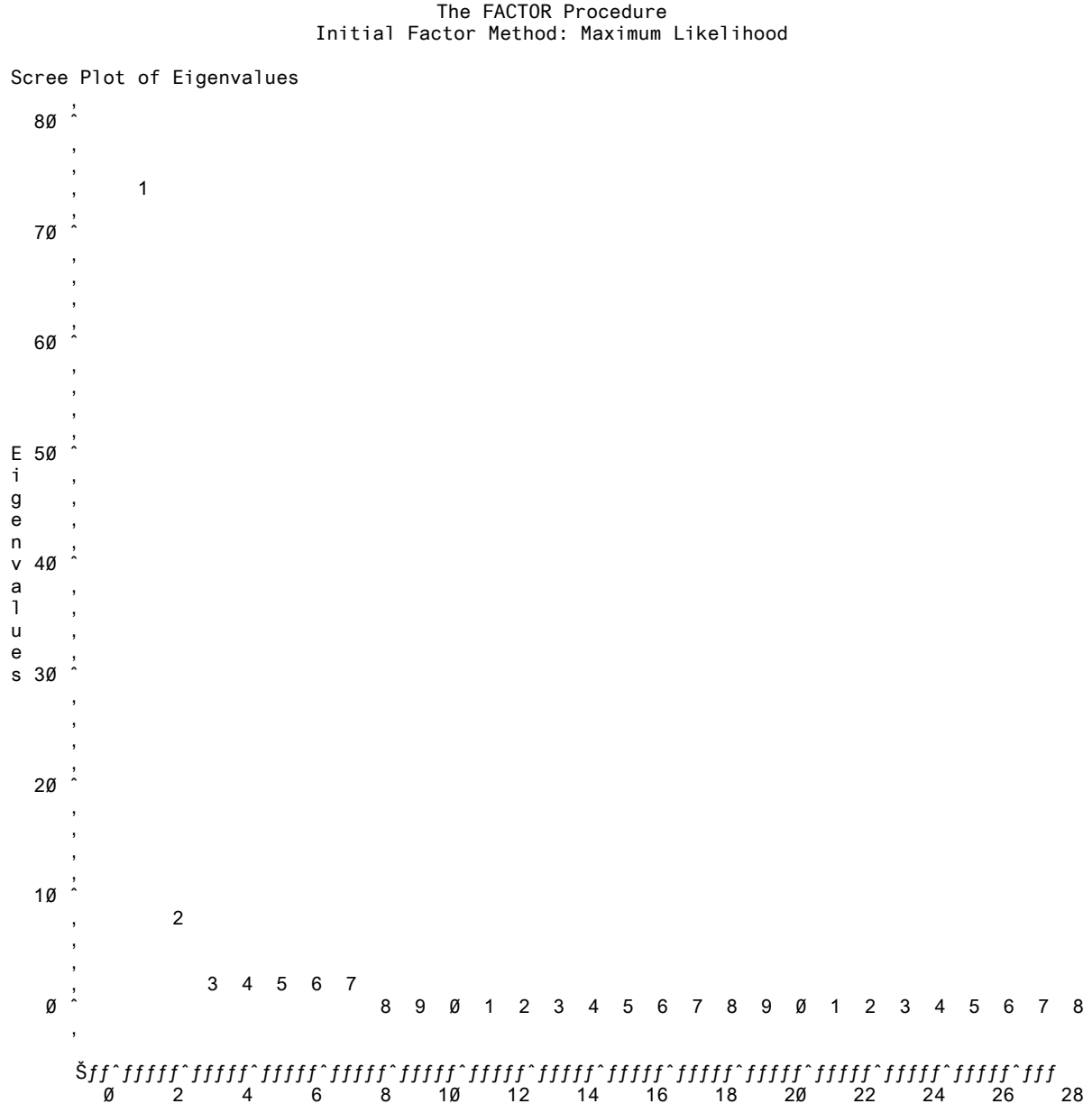
Group B Pre-test Responses



Group B Post-test Responses

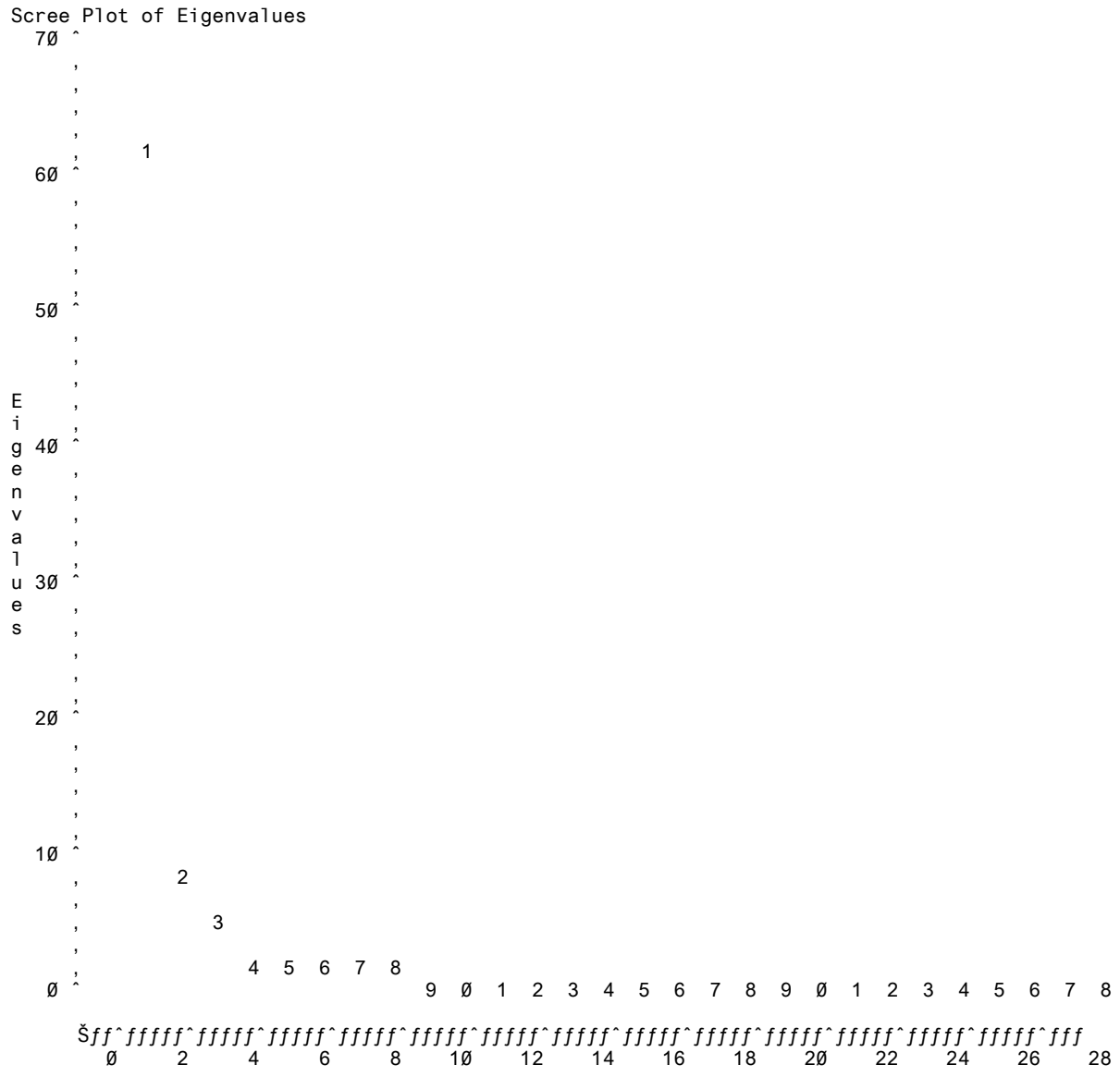


Group BC Pre-test Responses



Group BC Post-test Responses

The FACTOR Procedure
Initial Factor Method: Maximum Likelihood



Appendix H. Consent Form

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

An Empirical Study Comparing the Effect of Feedback, Training, and Coaching on Leadership Behavior Change

Nona E. Saling Paula Berardinelli

We are asking you to participate in a research study. The purpose of this study is to compare the effects of feedback, training, and coaching on leadership behavior change.

INFORMATION
If you agree to participate in this study, you will be asked to distribute feedback reports to up to fifteen superior(s), subordinates, and peers. You will receive a feedback report that you can compare to your first feedback report in order to assess how your leadership behavior has changed.

RISKS
The risks associated with receiving 360-degree feedback are minimal. If you have any questions about your feedback, please feel free to contact XXX at XXX, 800-XXX-XXXX, or your Human Resources representative.

BENEFITS
By participating in this study, you will be able to see the impact of your own efforts to enhance your leadership effectiveness.

CONFIDENTIALITY
The information in the study records will be kept strictly confidential. Data will be stored securely in a database with only a record number to identify you. No reference will be made in oral or written reports that could link you to the study.

COMPENSATION
For participating in this study, you will receive a free feedback report. If you withdraw from the study prior to its completion, or if fewer than three respondents complete and return feedback by the deadline date, you will not receive a feedback report.

CONTACT
If you have questions at any time about the study or the procedures, you may contact the researcher, Nona Saling, at (919) 419-7888 or at nsaling@mindspring.com. If you feel that you have not been treated according to the descriptions in this form or that your rights as a participant in research have been violated during the course of this project, you may contact Dr. Matthew Zingraff, Chair of the NCSU IRB for the Use of Human Subjects in Research Committee, Box 7514, NCSU Campus, (919) 513-1834; or Mr. Matthew Ronning, Assistant Vice Chancellor, Research Administration, Box 7514, NCSU Campus, (919) 513-2148.

PARTICIPATION
Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed, you will not receive a feedback report.

CONSENT
"I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may withdraw at any time."

Subject's signature_____ **Date** _____

Investigator's signature_____ **Date** _____

Please sign the two copies of this form included in your packet. Keep one for your records and return the other signed form to: **Research Study, XXX, XXX, XXX, XX XXXXX**